ROSA CRESPO RODRIGUEZ

THE PHYSICS OF THE PHASE TRANSITION IN CRYSTALLOGENESIS AND AMYLOID FIBRILLOGENESIS

Supervisor:

Doctor Pedro M. Martins, Departamento de Engenharia Química-Faculdade de Engenharia, Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto

Co-supervisors:

Doctor Fernando A. Rocha, Departamento de Engenharia Química-Faculdade de Engenharia, Universidade do Porto

Doctor Ana M. Damas, Instituto de Ciências Biomédicas Abel Salazar, Instituto de Biologia Molecular e Celular, Universidade do Porto

PORTO

January, 2016

A thesis submitted in part fulfilment of the requirement for an European Ph.D. degree in Chemical and Biological Engineering, University of Porto, Portugal

This work was financially supported by:

Project UID/EQU/00511/2013-LEPABE (Laboratory for Process Engineering, Environment, Biotechnology and Energy-EQU/00511)

FEDER funds through Programa Operacional Competitividade e Internacionalização - COMPETE2020

National funds through FCT (Fundação para a Ciência e a Tecnologia); Scholarship (SFRH / BD / 74729 / 2010).

Legal Precepts

According to the Decree Law No. 115/2013 published in *Diário da República*, 1st series nº 151 from August 7th 2013, the results of the following articles were used in this work:

"A generic Crystallization-Like Model that describes the kinetics of amyloid fibril formation" **Rosa Crespo**, Fernando Rocha, Ana M. Damas, Pedro M. Martins; Journal of Biological Chemistry, Aug. 2012, 287(36) 30585-94

"What can the kinetics of amyloid fibril formation tell about off-pathway aggregation?" **Rosa Crespo**, Eva Villar-Alvarez, Pablo Taboada, Fernando Rocha, Ana M. Damas, Pedro M. Martins; Journal of Biological Chemistry, Nov. 2015, 291, 2018-2032

"The availability of soluble protein is decreased in amyloidogenic media crowded with insoluble off-pathway aggreagtes" **Rosa Crespo**, Eva Villar-Alvarez, Pablo Taboada, Fernando Rocha, Ana M. Damas, Pedro M. Martins. Submitted, pendent of acceptance

Other publications:

"Potential use of ultrasound to promote protein crystallization" **Rosa Crespo**, Pedro M. Martins, Luís Gales, Fernando Rocha, Ana M. Damas; Journal of Applied Crystallography, Oct. 2010, 43, 1419-1425

"Small temperature oscillations promote protein crystallization" Cecilia Ferreira, **Rosa Crespo**, Pedro M. Martins, Luís Gales, Fernando Rocha, Ana M. Damas; CrystEngComm, Feb. 2011, xx, 1-7

"Fluorinated beta-sheet breaker peptides" Joana A. Loureiro, **Rosa Crespo**, Hans Börner, Pedro M. Martins, Fernando A. Rocha, Manuel Coelho, M. Carmo Pereira, Sandra Rocha; Journal of Materials Chemistry B, Jan. 2014, 2, 2259–2264

"Enzyme Kinetics: the whole picture reveals hidden meanings" Maria F. Pinto, Berta N. Estevinho, **Rosa Crespo**, Fernando A. Rocha, Ana M. Damas, Pedro M. Martins; The FEBS Journal, Mar. 2015, 282, 2309–2316

In compliance with the Decree Law, the author of this thesis states that intervened in the development and performance of the experimental work as well as in the interpretation and writing of the results published under the name **Rosa Crespo**.

“Satisfaction lies in the effort, not in the attainment, full effort is full victory”

-Mahatma Ghandi-

Acknowledgements

I would like to start by thanking my supervisor, Doctor Pedro Martins. Working with you has been definitely a remarkable experience in my life. I have learned so much with and from you during the last years. Thank you so much for your continued support, help and guidance, for always believing in me and for your sincere friendship.

I deeply thank my co-supervisors, Doctor Fernando Rocha and Doctor Ana M. Damas for their kindness, their support and for always helping and encouraging me. It has been my pleasure to have you both as my mentors.

It was a pleasure to work with Doctor Pablo Taboada and group. Thank you for receiving us so well in your lab and for all your help, for all the hours and effort to get such "good data" in this successful collaboration.

To Doctor Rosário Almeida, for her advice and help during the last period of this work. I also sincerely thank Frederico Silva and Joana Furtado from UP3 (IBMC) for helping me in the final adventure of protein production and purification.

I also would like to thank LEPABE-FEUP, ICBAS and IBMC, institutions where I developed all the experimental work involved in this project.

To my friends and colleagues and everyone that in a way or other made the difference in this journey. Thank you for your time, laughs, help, conversations, advice and good company.

Specially, and with all my heart, I thank my lovely family. It would not be possible to get here without you. Thank you for supporting and encouraging me so much along my entire life.

And finally, to the beautiful and colorful city of Porto…It became part of me!

Abstract

Amyloidosis covers a wide range of acquired or hereditary diseases frequently associated to protein misfolding. The soluble proteins are abnormally deposited extra or intracellularly as insoluble fibrillar aggregates, inducing toxicity and causing tissue and organ disruption and disease. More than 30 different unrelated proteins are known to form amyloid fibrils *in vivo* under specific conditions such as abnormally high concentration levels, genetic mutations or folding abnormalities. Amyloidosis can be systemic, such as Familial Amyloidotic Polyneuropathy, or located (organ-specific), such as Alzeimer´s or Prion disease, which are associated with protein aggregates in the brain.

In this work, fundamental principles of the physics of phase transitions are used to understand the process of amyloid fibril formation. Our strategy is to explore common aspects between amyloid fibrillogenesis and protein crystallization that have been identified by us during a pre-doctoral study. In both processes, an initial assembly of macromolecules in solution (nucleation step) is followed by the growth of the stable nuclei into insoluble supramolecular structures (growth step). Being energy-activated steps, nucleation and growth can be influenced by external factors such as, and most notably, the presence of additives. In fact, while crystallographers are challenged to obtain protein crystals for macromolecular structure determination, we aim at a better understanding of fibrillogenesis in order to be able to inhibit and not promote amyloid fibril formation. Our first goal was to characterize the mechanism of amyloid fibril formation by a rigorous, yet mathematically simple model that could be routinely used to describe protein aggregation data. Next, and following the wealth of published data on the clinical relevance of non-amyloidogenic aggregates, we extended our mathematical model to include a parallel step of off-pathway aggregation (OPA). This allowed us to identify a series of kinetic signatures revealing the quantitative importance of OPA relatively to amyloid fibrillization. We also found insoluble aggregates to produce solvation-in effects that decrease the thermodynamic concentration of protein in the system. The new theoretical tools were used on the characterization of amyloidogenesis inhibition, clearly distinguishing between nucleation/growth inhibitors (true inhibitors) and the pure thermodynamic effects caused by apparent inhibitors. An application example is included comprising preliminary screening tests of on/off-pathway inhibitors for α-synuclein (aSyn) aggregation, which is associated to Parkinson's disease.

This fundamental research is expected to contribute for the biophysical understanding of protein aggregation paving the way for development of innovative bioengineering tools for the therapy of neurodegenerative diseases.

Resumo

As amiloidoses abrangem uma ampla gama de doenças adquiridas ou hereditárias associadas frequentemente com alterações conformacionais de proteínas. As proteínas solúveis são anormalmente depositadas no interior ou exterior das células sob a forma de agregados fibrilares insolúveis, os quais podem ser tóxicos e causadores de perturbações em tecidos e órgãos. Mais de 30 proteínas diferentes são actualmente associadas à formação de fibras de amilóide *in vivo*, sendo para isso necessárias condições específicas tais como concentrações anormalmente elevadas, mutações genéticas ou alterações conformacionais induzidas. As amiloidoses podem ser sistémicas, como a Polineuropatia Amiloidótica Familiar, ou localizadas em órgãos específicos tais como a doença de Alzheimer ou as doenças priónicas (associadas com aggregados proteicos no cérebro).

Neste trabalho, os princípios fundamentais da física da transição de fase são usados para entender o processo de formação de fibras de amilóide. A nossa estratégia é explorar aspectos comuns entre a formação de fibras de amilóide e a cristalização de proteínas que identificamos durante os estudos pré-doutorais. Em ambos os processos, a associação inicial de um conjunto de macromoléculas em solução (passo de nucleação) é seguida pelo crescimento de núcleos estáveis que originam as estruturas supramoleculares insolúveis. Sendo os passos de nucleação e crescimento energeticamente ativados, são também susceptíveis de ser influenciados por factores externos, tais como, e em particular, a presença de aditivos. Enquanto que os cristalógrafos pretendem promover a formação de cristais de proteina para a determinação da estrutura macromolecular, a nossa intenção e como principal diferença, é que estudando em profundidade o processo de formação de fibras de amilóide queremos inibir a agregação. O nosso primeiro objectivo foi caracterizar o mecanismo de formação de fibras amilóide através de um modelo físico rigoroso, ainda que matematicamente simples de maneira a poder ser usado sistematicamente em ajustes de dados cinéticos de agregação de proteínas. A seguir, e com base num conjunto alargado de evidências experimentais apontando para a relevância clínica de agregados não-amiloidogénicos, optamos por alargar o modelo matemático de forma a contemplar um passo adicional de agregação *off-pathway* (OPA). Tal permitiu-nos identificar uma série de assinaturas cinéticas revelando quantitativamente a importância da OPA relativamente à formação de fibras de amilóide. Descobrimos também que os agregados insolúveis produzem efeitos denominados por *solvation-in,* os quais diminuem a concentração termodinâmica da

proteína no sistema. As ferramentas teóricas desenvolvidas foram usadas na caracterização da inibição de formação de fibras de amilóide, ajudando a distinguir entre a inibição dos passos de nucleação/crescimento (através de inibidores verdadeiros) e os efeitos termodinâmicos causados por inibidores aparentes. Um exemplo de aplicação é apresentado envolvendo testes preliminares de *screening* de inibidores *on/off-pathway* da agregação da proteína α-sinucleina (aSyn) associada à doença de Parkinson. Espera-se que esta investigação fundamental contribua para a compreensão biofísica do processo de formação de fibras de amilóide e para o desenvolvimento de novas ferramentas de base biotecnológica na terapia de doenças neurodegenerativas.

Resumen

Las amiloidosis cubren una amplia gama de enfermedades adquiridas o hereditarias causadas frecuentemente por alteraciones conformacionales de proteínas. Las proteínas solubles se depositan anormalmente extra o intracelularmente en forma de agregados fibrilares insolubles, los cuales pueden ser tóxicos y causar daños en tejidos y órganos. Más de 30 proteínas diferentes son actualmente asociadas a la formación de fibras de amiloide *in vivo*, bajo condiciones específicas tales como concentraciones anormalmente elevadas, mutaciones genéticas o alteraciones conformacionales inducidas. Las amiloidosis pueden ser sistémicas, como la Polineuropatia Amiloidótica Familiar, o localizadas en órganos específicos, tales como la enfermedad de Alzheimer o las enfermedades priónicas (asociadas com agregados proteicos en el cerebro).

En este trabajo, los principios fundamentales de la física de transición de fases se utilizan para entender el proceso de la formación de fibras de amiloide. Nuestra estrategia es explorar los aspectos comunes entre la formación de fibras de amiloide y la cristalización de proteínas que identificamos durante estudios predoctorales. En ambos procesos, la asociación inicial de un conjunto de macromoléculas en solución (paso de nucleación) es seguida por el crecimiento de núcleos estables que originan las estructuras supramoleculares insolubles. Siendo los pasos de nucleación y el crecimiento energéticamente activados, son también susceptibles a ser influenciados por factores externos, tales como, y en particular, la presencia de aditivos. En quanto los cristalógrafos pretenden promover la formación de cristales de proteína para la determinación de su estructura macromolecular, nuestra intención y como principal diferencia, estudiando en profundidad el proceso de formación de fibras de amiloide queremos inhibir y no promover la agregación. Nuestro primer objetivo fue caracterizar el mecanismo de formación de fibras de amilóide a través de un modelo físico riguroso, aunque matemáticamente sencillo de manera a poder ser utilizado sistemáticamente para el ajuste de datos cinéticos de agregación de proteínas. A continuación, e con base en un amplio conjunto de evidencias experimentales destacando la importancia clínica de agregados no amilodogénicos, optamos por alargar nuestro modelo matemático de forma a contemplar un paso adicional de agregación *off-pathway* (OPA). Esto nos permitió identificar una serie de características cinéticas revelando la importancia quantitativa de OPA relativamente a la formación de fibras de amiloide. También descubrimos que los agregados insolubles producen efectos denominados por *solvation-in*, disminuyendo la concentración

termodinámica de la proteína en el sistema. Las herramientas teóricas desarrolladas fueron aplicadas a la inhibición de formación de fibras de amiloide, ayudando a distinguir la inhibición entre los pasos de nucleación/crecimiento (a través de inhibidores verdaderos) y los efectos termodinámicos causados por inhibidores aparentes. Mostramos un ejemplo de aplicación implicando pruebas preliminares de *screening* de inhibidores *on/off-pathway* de agregación de la proteína α-sinucleína (aSyn), asociada a la enfermedad de Parkinson. Es de esperar que esta investigación fundamental contribuya a la comprensión biofísica del proceso de formación de fibras de amilóide y al desarrollo de nuevas herramientas con base biotecnológica para la terapia de enfermedades neurodegenerativas.

Table of Contents

List of Figures

List of Tables

Abbreviations

AA	Amyloid A Protein
AApoAI	Amyloid Apoliprotein AI
Aβ	Amyloid β peptide
Aβ2M	Amyloid β2-Microglobulin
AD	Alzheimer's disease
AFM	Atomic force microscopy
ALys	Amyloid Lysozyme
APro	Amyloid Prolactin
APrP	Amyloid Prion Protein
aSyn	Alpha-Synuclein
ATTR	Amyloid Transthyretin
BBB	Blood-brain barrier
CD	Circular Dichroism
CLM	Crystallization-Like Model
Da	Dalton
DLS	Dynamic Light Scattering
DMSO	Dimethylsulfoxide
DNA	Desoxiribonucleic acid
EDTA	Ethylene-diamine-tetraacetic acid
EGCG	Epigallocatechin-3-gallate
EM	Electron microscopy
F_∞	Fluorescence endpoint signals
FTIR	Fourier transform infrared spectroscopy
F-W	Finke and Watzky two-step model
HEWL	Hen egg-white Lysozyme

HPLC	High-performance liquid chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria broth
LCM-DIM	Laser confocal microscopy combined with differential interference contrast Microscopy
MS	Mass spectrometry
M-TTR	Monomeric TTR
NMR	Nuclear magnetic resonance
OPA	Off-pathway aggregation
PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline-Tween20
PD	Parkinson disease
PET	Position emission tomography
PIB	Pittsburgh Compound-B
QCM	Quartz crystal microbalance
R^2	Goodness of fit
SDS	Sodium dodecyl sulphate
SE	Standard error
SEC	Size exclusion chromatography
SNCA	Synuclein alpha [*Homo sapiens* (human)]
ThT	Thioflavin T
TIRFM	Total internal reflection fluorescence microscopy
UV	Ultraviolet
WT	Wild type

Chapter 1

General introduction

1.1 Common aspects in crystallogenesis and amyloid fibrillogenesis: the physics of the phase transition

Many analogies have been recognized between the physical processes of protein crystallization and amyloid fibrillization. As protein crystals, amyloid fibrils form in the supersaturated solutions of precursor proteins through a nucleation and growth mechanism characterized by a lag phase, an initial assembly of macromolecules in solution (nucleation step) followed by the growth of the stable nuclei into insoluble supramolecular structures (growth step).The lag phase can be by-passed via seeding (seed-dependent growth)[1].

Protein crystals and amyloid fibrils are well organized structures stabilized by common molecular interactions. The size and complexity of proteins is reflected in their ability to form many different intermolecular contacts, specific and nonspecific: hydrogen bonding, electrostatic and van der Waals interactions, as well as hydrophobic interactions and volume excluded effects. All those interactions involved in the assembly of macromolecules are affected by temperature, pH, protein concentration, ionic strength of the solution and concentration of specific additives[2–4].

Being energy-activated steps, nucleation and growth can be influenced by external factors. Many recent studies showed that the nucleation step in protein aggregation can be accelerated by stirring and by using ultrasonication, previously recognized as a characteristic of the crystallization process[3–5]. This common dependency on ultrasonication can be explained for both of the processes by a common phase diagram (Figure 1.1)[4].

In both processes, the supersaturated protein under study (or peptide) seems to adopt a water-excluded ordered assembly. In those cases, supersaturation refers to a metastable state in which the solution keeps kinetically stable although the concentration of dissolved solutes is larger than the thermodynamic solubility. As a consequence of a high free energy barrier, in the metastable state proteins hold soluble for an extended period[3,5].

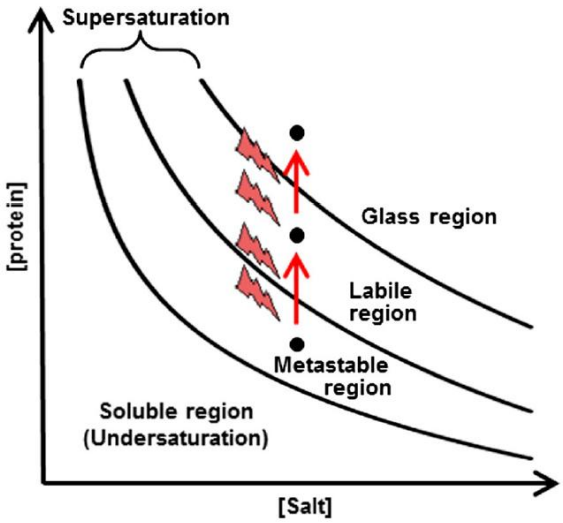


Figure 1.1. Common phase diagram for protein crystallization and fibrillization addressing the acceleration of the nucleation step by ultrasonic irradiation, indicated by the red narrows. Adapted from Kitayama et al. (2013)[4].

It does not seem to be a coincidence that by the time we proposed our two-step mathematical model for describing amyloid aggregation kinetics in 2012[6], and in which supersaturation is considered as the major driving force for protein fibrillization (common aspect with protein crystallization and based on our previous work[5,7]), many other studies confirming this evidence have appeared[1,3,4,8–13].

1.2. Amyloid

Used in botany, the term 'amyloid' was first applied by Rudolf Virchow in 1854 to name a human alteration, cerebral corpora amylacea[14]. Presenting pale blue stained when treated with iodine and violet after the addition of sulfuric acid, Virchow declared it cellulose and gave it the name 'amyloid'. Amyloid comes from the Latin term amylum and the Greek amylon[15]. This iodine staining characteristic was also shown in Virchow's analysis of spleen, liver and kidney deposits, in cases of supposed amyloid A (AA) amyloidosis due to chronic infectious diseases. Assuming the carbohydrate rich nature of those tissue accumulations, they were named as amyloid deposits[16].

The absence of carbohydrate accumulation and high content in nitrogen was demonstrated in 1859 by Friedreich and Kekule, who established the proteaceous rich composition of amyloid deposits[15]. Even that, the nature of amyloid proteins remained

unclear for a long time. During the 20th century, the development of experimental techniques allowed to further investigations on amyloid and made possible a more precise description[17].

At the beginning, amyloid was refered to the soluble protein abnormally deposited extracellularly as insoluble highly ordered fibrillar aggregates, inducing toxicity and causing tissue and organ disruption and disease[18]. Classically recognized by their affinity to Congo red showing green birefringence under polarized light, those fibrillar aggregates were called amyloid fibrils. Later, intracellular deposition of amyloid fibrils was founded, missing the affinity to Congo red in most cases[19].

In the last decades, a key discovery happened when amyloid proteins were found to form amyloid fibrils in vitro. Many other proteins with no pathological role in disease and non-natural polypeptides have also shown to form amyloid fibrils in vitro under appropriate conditions[20,21]. All them presenting binding affinity to Congo red and other amyloid specific dyes. In those cases, the term 'amyloid-like' is frequently applied and 'amylog' was suggested when there is no disease-association[14]. From a more biophysical point of view, it is generally accepted to name as 'amyloid' any of those specific fibrillar structures independently of their nature. In this context, amyloid fibrils are defined as self-assembled, filamentous protein or peptide aggregates with cross-β-sheet structures running perpendicular to the long fibril axis[22,23], meaning they contain ribbon-like β-sheets in which the β-strands run perpendicular to the long fibril axis and the inter-strand hydrogen bonds run parallel to the long fibril axis[24].

Nowadays, amyloid aggregation is implicated in a wide range of scientific disciplines, highlighting its recent interest in biotechnology. However, extensive research has primarily been driven by its association with human disease[25–27]. The total number of proteins in the body can be estimated between 300,000 and 1 million, from the 30,000-40,000 protein-coding genes that appear to be in the human genome[28] and of which more than 30 are known to aggregate forming disease-associated amyloid fibrils[29]. Around 100 different human pathological conditions are recognized to involve amyloid fibril aggregation (Table 1.1)[23,30].

Table 1.1. Selected examples of polypeptides known to form disease-associated amyloid deposits[14,23].

Precusor polypeptide	**Amyloid classification**	**Associated syndrome/disease**
Aβ peptide	Aβ	Alzheimer's disease
β_2-Microglobulin	A β_2M	Dialysis related amyloidosis
Prion protein	APrP	Spongiform encephalopathies
Prolactin	APro	Prolactinomas
Transthyretin	ATTR	Familial amyloid polyneuropathy
Lysozyme	ALys	Hereditary non-neuropathic systemic amyloidosis
Serum amyloid A protein	AA	Rheumatoid arthritis
Alpha-synuclein	none	Parkinson's disease
Huntingtin	none	Huntington's disease
Apolipoprotein AI	AApoAI	Atherosclerosis

The fact of a wide range of different proteins/peptides forming amyloid fibrils in vitro has made easier further studies of amyloid aggregation. In the last decades, experimental and theoretical studies have been developed, disclosing many key features and raising a number of other new important issues that at the moment remain unclear[31].

1.3 Following amyloid aggregation

A great deal of effort has been put into developing biophysical techniques for describing amyloid aggregation in the last decades. The identification of amyloid fibrils structure and oligomeric species as well as the pathway of association are crucial for a complete description of the process[25]. Structural information can be obtained by circular dichroism (CD), Fourier transform infrared spectroscopy (FTIR), size exclusion

chromatography (SEC), nuclear magnetic resonance (NMR) imaging[32] and/or mass spectrometry[33]. Electron microscopy (EM), atomic force microscopy (AFM) and X-ray diffraction are frequently used to investigate intermediate oligomeric species and final products of aggregation[34].

Elucidating the molecular mechanism involved in amyloid aggregation has been one of the most pursued issues in order to develop anti-amyloid therapeutics. Kinetic and thermodynamic measurements of amyloid fibril formation are essential to achieve this goal. Despite the limitations of the classical methods used to follow amyloid aggregation in vitro, kinetic and thermodynamic data can be measured by different techniques including light scattering[35], turbidity/absorbance[36], fluorescence spectroscopy (intrinsic fluorophores and extrinsic dyes) and/or high-performance liquid chromatography (HPLC)[37]. More recently, other techniques have been developed to follow fibril growth as AFM[38,39], total internal reflection fluorescence microscopy (TIRFM) and quartz crystal microbalance (QCM)[25,40]. Laser confocal microscopy combined with differential interference contrast microscopy (LCM-DIM) is an advanced microscopy technique that was originally developed to measure 2D nucleation kinetics of protein crystals[41], and that represents a promising technique to follow the different phases during amyloid fibrillation. However, each method commonly reports a specific parameter of the aggregation system, such as mass of aggregates, monomer concentration, average filament length, etc. Complementary methods are often required for consistent analysis[25,26,34,42,43].

Covering the most commonly used methods for studying the whole amyloid aspects, Li et al. designed a detailed and consistent scheme (Table 1.2), useful to get a general idea of the current analyses outlook[38].

Table 1.2. Most common biophysical methods used to study amyloid protein structure,amyloid aggregation, and the different assemblies involved in the process. Adapted from Li et al.[38]

<table>
<tr><th></th><th>Monomer</th><th>Oligomer</th><th>Protofibril</th><th>Fibril</th><th>Amyloid (in vivo)</th></tr>
<tr><td rowspan="3">Atomic Structure</td><td colspan="2">Solution-state NMR</td><td></td><td>Solid-state</td><td></td></tr>
<tr><td></td><td></td><td></td><td>X-Ray</td><td></td></tr>
<tr><td colspan="4">X-Ray absorption</td><td></td></tr>
<tr><td rowspan="8">Secondary Structure</td><td colspan="3">Circular dichroism spectroscopy</td><td></td><td></td></tr>
<tr><td colspan="4">Fourier transform infrared spectroscopy</td><td></td></tr>
<tr><td></td><td></td><td colspan="2">X-Ray fiber diffraction</td><td></td></tr>
<tr><td></td><td></td><td colspan="2">Neutron scattering</td><td></td></tr>
<tr><td></td><td></td><td colspan="3">Congo red binding/ Thioflavin S fluorescence</td></tr>
<tr><td></td><td></td><td></td><td colspan="2">Birefringence</td></tr>
<tr><td></td><td></td><td colspan="2">Thioflavin T fluorescence</td><td>PIB/FDDNP/SB13/BF227</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="4">Morphology</td><td></td><td colspan="3">Transmission electron microscopy</td><td></td></tr>
<tr><td></td><td colspan="3">Scanning transmission electron</td><td></td></tr>
<tr><td></td><td colspan="3">Scanning tunneling microscopy</td><td></td></tr>
<tr><td></td><td colspan="3">Atomic force microscopy</td><td></td></tr>
<tr><td rowspan="4">Tertiary/quaternary Structure</td><td colspan="4">Electron spin resonance</td><td></td></tr>
<tr><td colspan="4">Hydrogen-Deuteriom exchange</td><td></td></tr>
<tr><td colspan="4">Limited proteolysis</td><td></td></tr>
<tr><td colspan="4">Fluorescence methodologies</td><td></td></tr>
<tr><td rowspan="5">Assembly size/ Size distribution</td><td colspan="2">Gel electrophoresis</td><td></td><td></td><td></td></tr>
<tr><td colspan="2">PICUP</td><td></td><td></td><td></td></tr>
<tr><td colspan="3">Size-exclusion chromatography</td><td></td><td></td></tr>
<tr><td colspan="3">Analytical ultracentrifugation</td><td></td><td></td></tr>
<tr><td colspan="3">Dynamic light scattering</td><td></td><td></td></tr>
<tr><td colspan="2">IMS-MS</td><td></td><td></td><td></td></tr>
</table>

Thioflavin-T (ThT) fluorescence spectroscopy is perhaps the most widely adopted technique among the entire history of the in vitro amyloid aggregation. First described in 1959 by Vassar and Culling, the dye ThT was later characterized by Naiki et al. and LeVine who proof its specific interaction with the cross-β-sheet structure common to amyloid fibrils[44,45]. ThT molecule, formed by a benzylamine and a benzathiole ring sharing a carbon-carbon bond that rotates freely in solution, displays a dramatic shift of the

excitation and emission maximum (~440 and 482 nm respectively) upon binding of fibrils. ThT fluorescence originates only from the specific dye when the binding takes place by blocking its rings rotation[44,45]. These characteristics make from ThT a very sensitive and efficient reporter for real time fibrillization in solution, one of the main advantages over other dyes as Congo red[46]. The staining process with Congo redresults laborious, polarized light microscopy is required and the accurate interpretation of the apple green birefringence signal can be unclear[47]

Recent advances have been developed enabling direct visualization of amyloid plaques in the living human brain using position emission tomography (PET) and radioactive tracers with higth affinity towards amyloid with the ability to cross the blood-brain barrier (BBB)[48] . One of the most successful radiotracers consist in a ThT derivative called Pittsburgh Compound-B (PIB)[49]. Much research in neurodegeneration is being developed for finding new amyloid ligands enabling to further studies in the in vivo amyloid panorama[50].

1.4 Fundamental models for amyloid aggregation

Different approaches can be found in the literature to determine the kinetics and the mechanism of amyloid aggregation. The development of biophysical techniques that permit to follow the kinetics and thermodynamics of the amyloid aggregation process has made possible the appearance of fundamental models based on experimental data. Those models are our main focus in this work, but other approaches are also available, including empirical equations. The logistic equation has been extensively used to fit amyloid aggregation kinetic data. Although offering a convenient and easy option, there is no physical meaning behind the classic empirical equations, which cannot provide mechanistic information about the process[34,51].

Kinetic and thermodynamic descriptions of protein aggregation were recently reviewed by Gillam and MacPhee[34] and previously, but also of special interest, by Morris, Watzky and Fink[51]. A complete and detailed overview of the available fundamental models historically classified as well as by their scientific base is provided in these reviews.

The first key kinetic model for protein aggregation was proposed in 1959 by Oosawa et al., describing the total concentration of monomers involved in the formation of

helical polymers during actin fibrillization. This work and following contributions are based in a mechanism of primary nucleation (homogeneous nucleation) with subsequent monomer addition for polymerization. However, evidences of secondary growth processes affecting fibril formation were recognized. Ferrone et al[52] in 1980 made a contribution to this type of model by introducing the concept of monomer-dependent secondary nucleation (heterogeneous nucleation). In contrast to homogeneous nucleation, heterogeneous nucleation occurs on the surface of existing fibrils and thus depend on the concentrations of both monomeric proteins and fibrils[53]. Oosawa et al. and Wegner and Savko introduce further steps of secondary growth processes as filament fragmentation and association for breakable amyloid filaments[51,54]. The latest contributions of this type of mechanism, and both of special interest are Ferrone's (1999)[55] and Knowles et al. (2009)[56].

In 1997, the Ockham's razor minimalistic 2-step model was proposed by Finke and Watky for transition-metal nanocluster formation[57]. Later on, it was successfully applied in fitting a wide range of aggregating proteins, relevant in neurodegenerative disorders[51]. Based on a slow nucleation followed by a fast autocatalytic growth with rate constants k_1 and k_2 respectively, the Finke and Watzky two-step model (F-W) is represented by the following scheme:

$$Slow\ Nucleation: A \xrightarrow{k_1} B$$

$$Growth: A + B \xrightarrow{k_2} 2B$$

Scheme 1. The Finke and Watzky mechanism[34] scheme representation, where A represents the precatalytic form of the monomer and B corresponds to any catalytic form of the aggregated protein past the critical nucleus size.

Rate equations to fit amyloid aggregation kinetic data are obtained by direct integration of this simple model. Equivalent mechanism were use to fit protein fibrillization by Saitô et al. in 2000, presenting a fractional form of the F-W model, and Murphy et al. in 2006[34,51].

Although most of the available models fit successfully the classical sigmoidal kinetic curve of protein fibrillization in which an initial lag phase is followed by exponential growth and a final plateau level of aggregation, there are many other common behaviors that remain unclear and mechanisms fail in their explanation. The growth of amyloid fibrils is a complex process comprising many different parallel steps and it becomes complicated to obtain a universal analytical solution when it is not clear which is

the dominant step of the process[34,51]. Evidences of the limitations of the Oosawa and Ockham's razor type models are openly discussed as well as the inconsistencies with experimental data[34].

In 2012 we made a contribution by proposing a new mathematical model that describes successfully the kinetics of amyloid fibril formation for a wide range of different conditions and different amyloidogenic proteins, the Crystallization-Like Model (CLM)[6]. Along this work, the CLM will be presented and explained in detail. Further comparisons with the available models of the literature and analysis will be openly discussed.

1.5 Therapeutic strategies targeting protein aggregation

Over the last decades, remarkable progress has been made towards the development of therapeutic strategies targeting amyloidogenic diseases as a consequence of the important features that have been disclosed by the extensive research in protein aggregation[29,58]. The most promising therapeutic strategies under preclinical, clinical trials or clinical practice at the moment are based on the reduction of the amyloidogenic protein production, its stabilization, regulation of the proteostasis network and direct inhibition of the self-assembly process by amyloid remodeling and/or removal[29].

The combination of preventive strategies, acting before protein aggregation takes place or at an early beginning of the process, interfering with intermediate aggregates with strategies removing amyloid fibrils will be an encouraging option depending on the nature of the protein and the mechanism involved in its fibrillization. However, the mechanistic connection between the process of protein aggregation and tissue degeneration remains unclear, as well as which of the aggregate types involved in the process are toxic and their mechanism of cytotoxicity. A more detailed understanding of those key factors is needed to develop future therapies achiving successful results[29].

Small molecule and macromolecular approaches are currently the most promising options for the intervention of amyloid-associated disorders. Other approaches are adopted under specific conditions: in the case of amiloidogenic transthyretin (ATTR), mainly produced in the liver and causing TTR amyloidosis/hereditary fibrinogen amyloidosis, protein reduction is achieved as a standard of care by liver transplantation[29].

In this work, our focus will be in inhibiting protein aggregation with small molecules identified via high-throughput screening. This small molecule inhibitors approach was initially based on early findings that demonstrated that small aromatic molecules such as Congo red and Thioflavin T interact specifically with amyloid fibrils and inhibit their formation[59,60]. Recent effort has focused on small molecule inhibitors of amyloid fibril formation and as a consequence, many reports have been appearing during the past few years[60].

Chapter 2

Protein aggregation described by a two-parameter model

A generic Crystallization-Like Model that describes the kinetics of amyloid fibril formation

Rosa Crespo[1], Fernando A. Rocha[1], Ana M. Damas[2,3], and Pedro M. Martins[1,2]

[1]From LEPAE, Departamento de Engenharia Química, Faculdade de Engenharia da Universidade do Porto, Portugal

[2]ICBAS – Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto, Portugal

[3]IBMC – Instituto de Biologia Molecular e Celular, Porto, Portugal

To whom correspondence should be addressed: Pedro M. Martins, ICBAS – Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Rua de Jorge Viterbo Ferreira n.º 228, 4050-313 Porto, Portugal, Tel.: (+351) 222062287; Fax: (+351) 222062232; E-Mail: pmmartins@icbas.up.pt

Keywords: Amyloid, biophysics, mathematical modeling

Summary

Associated to neurodegenerative disorders such as Alzheimer's, Parkinson's, or Prion diseases, the conversion of soluble proteins into amyloid fibrils remains poorly understood. Extensive "in vitro" measurements of protein aggregation kinetics have been reported but no consensus mechanism has emerged until now. This contribution aims at overcoming this gap by proposing a theoretically consistent Crystallization-Like Model (CLM) that is able to describe the classic types of amyloid fibrillization kinetics identified in our literature survey. Amyloid conversion represented as a function of time is shown to follow different curve shapes, ranging from sigmoidal to hyperbolic, according to the relative importance of the nucleation and growth steps. Using the CLM, apparently unrelated data is deconvoluted into generic mechanistic information integrating the combined influence of seeding, nucleation, growth and fibril breakage events. It is notable that this complex assembly of interdependent events is ultimately reduced to a mathematically simple model, whose two parameters can be determined by little more than visual inspection. The good fitting results obtained for all cases confirm the CLM as a good approximation to the generalized underlying principle governing amyloid fibrillization. A perspective is presented on possible applications of the CLM during the development of new targets for amyloid disease therapeutics.

2.1 Introduction

Amyloid fibrils are insoluble ordered structures sharing a common cross β-sheet conformation and formed by misassembly of soluble proteins and peptides[61–63]. Since these protein aggregates are associated with numerous neurodegenerative diseases, a great deal of effort has been put into understanding the mechanisms of amyloid fibril formation[51,64,65]. Several biophysical methods have been developed to measure the kinetics of protein aggregation *in vitro*[25,42,43,56,66]. Thioflavin-T binding fluorescence[67] and turbidity measurements[36]are among the most widely adopted techniques. Despite the limitations of the classical methods[26], they provide quantitative information about the mass increase of amyloid fibrils with time. Typically, the progress of fibrillization is expressed as the normalized fraction of amyloid protein converted into fibrils (α) – on a mass basis, α corresponds tothe quotient of the fibril mass increase Δm at a given instant divided by the total mass of fibrils formed at the end of the assay Δm_T:

$$\alpha = \frac{\Delta m}{\Delta m_T} \qquad (2.1)$$

The time-course results represented in Figure 2.1 exemplify well two types of kinetics usually observed during amyloid fibril formation. The sigmoidal $\alpha(t)$ trend obtained for β_2-microglobulin (β_2m)[68] (open symbols) is characteristic of nucleation-dependent polymerization: an initial lag phase is followed by a phase of rapid growth and then by a stationary phase[40,69]. The hyperbolic shape obtained for transthyretin (full symbols) is more frequently found during seeded aggregation, i.e., when an aliquot of solution containing preformed fibrils is added to the amyloidogenic solution in order to bypass the thermodynamically unfavorable nucleation step[25,40,70]. Nevertheless, as shown by this example, the absence of a lag phase is also reported for unseeded reactions[70,71].

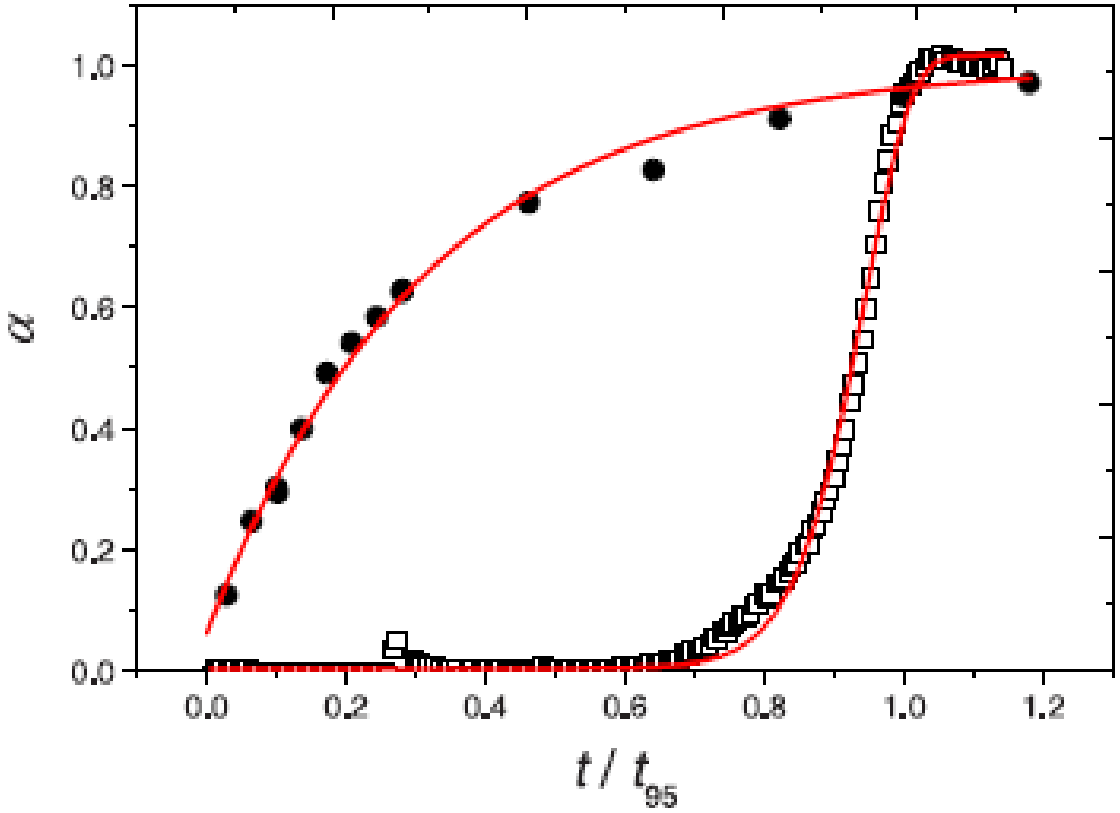


Figure 2.1. Time dependent normalized fluorescence signal representing typical sigmoidal (open symbols) and hyperbolic (full symbols) aggregation kinetics of amyloid proteins during amyloidogenesis. For visual clarity, trend lines are represented and the time axis is normalized by the time for 95% conversion (t_{95}) for each case. Open symbols correspond to the experimental data obtained by Xue et al.[68] using 61 μM of β_2m; t_{95}=15.6 hr. Full symbols represent continuous fluorescence data from Hurshman et al.[70] obtained for 0.15 mg/mL of monomeric transthyretin (M-TTR), which we digitized into periodic data and normalized by the endpoint signal; t_{95} = 23.4 s.

Between the classical sigmoidal kinetics, showing an inflection point at $\alpha_i = 0.5$, and the hyperbolic/seeded polymerization kinetics, showing no evident inflection point, there are a number of intermediate possibilities that have been comprehensively reviewed by Finke and co-workers[51,72]. As discussed in those reviews, the meaning of the different shapes has been interpreted over the last 50 years according to several thermodynamic and kinetic mechanisms. Given the complexity and variability of the protein aggregation models, a modular approach was recently proposed in order to systematically identify the mechanism that best describes nucleation (or prepolymerization), growth (or polymerization) and fragmentation steps[68]. The available fundamental models are in general highly specific for the amyloid polypeptide under study. Models containing a high number of parameters cannot fit to experimental data in a unique way, while some of the

theoretical elementary steps refer to time frames that are experimentally inaccessible. Overparameterization makes it hard to validate different parts of mechanistically elaborated model[73], being recognized as a problem in protein aggregation modeling[74]. In the pursuit of a consensual, two-parameter model for describing protein aggregation kinetics, the 1997 Finke-Watzky (F–W) mechanism for transition-metal nanocluster formation[57] was successfully applied to a wide range of kinetic data[51,75]. In the formulation of the F–W model, slow continuous nucleation is followed by autocatalytic surface growth so that the overall rate of protein concentration decrease is given by

$$\frac{-\mathrm{d}C}{\mathrm{d}t} = k_1 C + k_2 C\left(C_0 - C\right) \qquad (2.2)$$

with k_1 and k_2 being, respectively, the nucleation and growth average rate constants. For an initial protein concentration C_0, the solution of the ordinary differential equation can be expressed as

$$\frac{C_0 - C}{C_0} = 1 - \frac{1 + \frac{k_1}{k_2 C_0}}{1 + \frac{k_1}{k_2 C_0}\exp\left\{\left(1 + \frac{k_1}{k_2 C_0}\right) k_2 C_0 t\right\}} \qquad (2.3)$$

Note that the difference $(C_0 - C)$ corresponds, by mass balance, to the instantaneous concentration of protein aggregates[75]. The F–W equation is interconvertible by algebraic manipulation into the Saitô's equation[32] and into Fernández et al. equation[76], which however depart from different mechanistic assumptions (see comparative study below). When the nucleation rates are very low ($k_1 \to 0$), eq. 2.2 reduces to the differential form of the logistic function, which is commonly used to fit sigmoidal aggregation data[51,69]. Ferrone's equation is another example of a two-parameter protein aggregation model available in literature that predicts the monomer concentration to decrease with the square of time[55]. While avoiding overparameterization, these models are very useful from a practical point of view, even though some simplifying hypothesis adopted during their derivation may lack solid biophysical basis. For example, Saitô and co-workers, and Fernández et al. describe elementary kinetic steps by reaction-like rate equations that are first-order in relation to the fractional conversion of monomers into fibrils expressed as α and/or $(1–\alpha)$[32,76]. In an alternative formalism, Finke and Watzky consider the fibrillar state as the autocatalytic, polymeric form of the protein and express the elementary rate laws as a function of the protein concentrations in the solubilized and aggregated forms[51,72]. Although theoretically more appropriate, the F–W formalism results inthe paradox that,

according to eq. 2.2, all the dissolved protein is expected to convert into fibrils, i.e., the expected steady-state concentration is $C_\infty = 0$. This is not possible in the light of the general phase equilibrium condition and is not supported by measurements of dissolved protein concentration after long incubation times[70,77,78]. Finally, the applicability of Ferrone's quadratic equation is inherently limited to the early data points corresponding to ~10–20 % of total monomer loss[74].

Kinetic modeling of amyloid fibrillization reactions has also been performed using empirical equations such as linear and exponential decay functions[79]. As recently pointed out by Auer and Kashchiev while discussing the applicability of the Avrami equation[80], it does not seem coincidence that such mathematically simple models are able to describe the $\alpha(t)$ dependence for a wide range of polypeptides and amyloidogenic conditions. With the present contribution, we seek the general principle that seems to govern the kinetics of protein aggregation. We intend to do it in a theoretically consistent way that does not compromise the final simplicity of the model nor its quantitative usefulness. Infinite number of parameter combinations producing equally good agreement with the data should be avoided; that is to say, all the dynamic state variables shall be condensed in a two-parameter model that can be uniquely fit to protein aggregation kinetics. Then, the deepest meaning of each constant should be possible to be determined by using different types of experimental data. For this mechanistic refinement to be possible, oversimplified hypothesis such as those identified in our literature survey should also be avoided during the derivation of the model.

2.2 Crystallization-Like Model (CLM)

Resemblances between amyloid fibril formation and protein crystallization have long been recognized[40,81]. Both processes involve the thermodynamic equilibrium between phases, an initial assembly of macromolecules into stable nuclei (nucleation step), and the subsequent formation of supramolecular structures by successive addition of growth units (growth or elongation step). We propose now to quantitatively describe the aggregation kinetics ofamyloid proteins using fundamental principles that are familiar to Crystal Growth scientists.

2.2.1 Thermodynamic driving force

We will start by defining the thermodynamic driving force for amyloid fibril formation as the variation in chemical potential $\Delta\mu$ occurring when protein molecules in a supersaturated amyloid solution at temperature *T* are transferred to a fibrillar state:

$$\Delta\mu = kT \ln \frac{a}{a^*} \qquad (2.4)$$

where*k* is the Boltzmann constant, and *a* and *a** are the activities of supersaturated and saturated amyloid solutions. This is analogous to the definition of supersaturation adopted in crystallization from solution[82]. Accurate activity coefficients are difficult to obtain for highly non-ideal concentrated solutions. For this reason, and for mathematical simplicity, we will define an approximate amyloid supersaturation σ as a function of protein concentration *C* and protein solubility *C**:

$$\sigma = \frac{C - C^*}{C^*} \qquad (2.5)$$

The protein solubility corresponds to the concentration of dissolved protein that equilibrates the chemical potential of the insoluble fibrillar phase. In principle, *C** is independent of the initial protein concentration and can be obtained from the remaining concentration of the polypeptide in solution after long reaction times[70]. Nevertheless, when measuring protein aggregation in the presence of high macromolecular content ("macromolecular crowding"), volume-excluding effects may lead to different protein concentrations in equilibrium[83]. We will return to this issue while discussing possible deviations from the CLM. The total amount of aggregates produced per unit of volume of solution $\Delta m_T / V$ is given by the difference between the mass concentrations C_0 and *C**. Equation 2.5 can thus be re-written to express amyloid supersaturation as a function of the fraction of protein converted into amyloid aggregates (α) and of the initial supersaturation (σ_0):

$$\sigma = \frac{(C_0 - C^*) - (C_0 - C)}{C^*} = \\ = \frac{\Delta m_T}{VC^*}(1-\alpha) = \sigma_0 (1-\alpha) \qquad (2.6)$$

2.2.2 Nucleation step

The mechanism leading to the formation of amyloid nuclei generally includes conformational changes of the native state[84] and different intermediate structures such as polymorphous and oligomeric aggregates[85,86], whose nature and importance varies from

protein to protein[51,68]. While all of the transitions involve energy barriers of different magnitudes, we will consider that the rate limiting step is the formation of a critical-sized amyloid nucleus. This postulation finds support in the well-known fact that the addition of preformed fibrils to amyloidogenic solutions (seeding) completely eliminates the lag phase and induces immediate fibril formation. Classical Nucleation Theory[87,88], based on which numerous phase transition phenomena occurring in nature and technology have been explained[89], will be used to estimate the frequency at which new fibrils are created. It is admitted that fluctuations in phase density give rise to the appearance of embryonary formations or nuclei that reduce the bulk free energy (due to the variation in chemical potential $\Delta\mu$) and increase the surface free energy (due to the creation of a fibril/solution interface). Above a critical nucleus size the process spontaneously evolves in the direction of amyloid fibril formation. Recently[88], the following expression was proposed for the critical number of monomeric peptides constituting the amyloid nucleus:

$$n^* = \frac{\nu}{\Delta\mu^2} + 1 \qquad (2.7)$$

whereν is a constant accounting for the dimensions and interfacial energies of the nanosized fibril. In the same work, the nucleation rate J was expressed as a complex function of the thermodynamic driving force for amyloid fibril formation[88]. We will adopt a simplified version of this relationship that takes into account the very high supersaturation levels associated to amyloidogenic conditions[70,77,78]. For C/C^* ratios much higher than 10, the equilibrium concentration of fibril nuclei is expected to increase linearly with C, greatly simplifying the expression of J to a second-order rate equation:

$$J \cong A\sigma^2 \qquad (2.8)$$

with A being a σ-independent kinetic factor expressed in units of frequency per solution volume.

2.2.3 Growth step

During this step, preformed fibrils grow by successive addition of new structural units. Due to the existence of a preliminary fibril template, the growth (or elongation) step is thermodynamically favored over nucleation; although both processes co-exist during the early phases of incubation[69], fibril elongation is expected to become increasingly predominant[90]. Advanced microscopy techniques show indefinite evidences of unidirectional[91–94] and bidirectional amyloid fibril growth[39,95], reportedly involving

intertwining of amyloid filaments and protofibrils[96]. Whatever the mechanisms of assembly admitted to occur, phase transition during this step is mostly determined by a secondary nucleation process taking place on the surface of the protein aggregates[53,94]. Therefore, the elongation rate G is proportional to the frequency at which protein molecules leave the solution phase and reach the fibril surface. As previously stated by us for the case of crystal growth from solution[97–99], the attachment probability linearly increases with (i) supersaturation σ, which drives the diffusion of molecules toward the interface, and with (ii) the number of energetically favorable sites for integration. The later quantity was found to be proportional to crystal size during crystallization[97], and is here assumed to be proportional to the length of amyloid fibrils L or, more specifically, to the number of β-sheets constituting the solid phase. We depart from a non-active binding mechanism, where the lateral binding of monomers is seen as a dynamic reservoir for their subsequent incorporation at both ends of the fibrils[100,101]– Figure 2.2. We further assume that surface adsorption is a reversible process followed by a fast surface diffusion step, which is not rate-limiting while monomers are being recruited into the active sites. The simplified expression for the rate of fibril elongation is thus given by

$$G = \frac{\mathrm{d}L}{\mathrm{d}t} = k_g \sigma L \qquad (2.9)$$

where k_g is a growth rate constant accounting for the influx of amyloidogenic protein molecules arriving to the fibril ends, and undergoing conformational changes upon integration. For constant cross-section area, the mass of fibrils linearly increases with L. Therefore, size-dependent growth provides a phenomenological explanation for the autocatalytic nature of amyloid protein aggregation[69,102], and for the semi-empirical evidences that the elongation rate is proportional to the mass of fibrils, directly as m-dependent growth[74], or indirectly as α-dependent growth[32] and (C_0–C)-dependent growth[51,75]. A more subtle consequence arising from eq. 2.9 is that, unless specified otherwise by complementary rate laws, overall elongation rates are affected by the total mass of fibrils present in solution but not by their number. We will return to this point during the derivation of the mathematical model and during the discussion of its limitations.

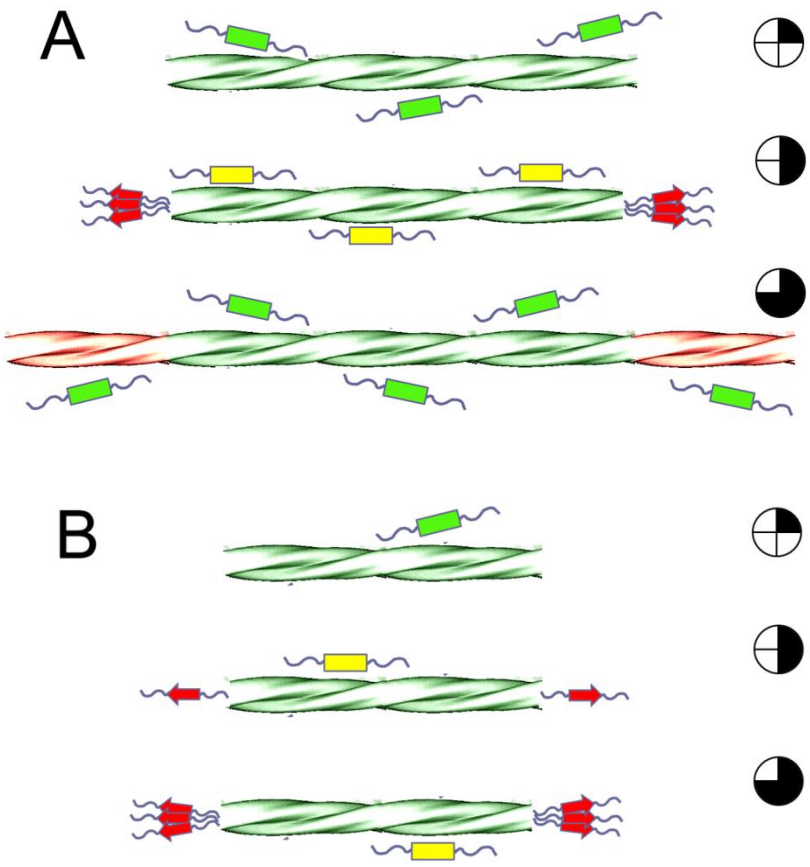


Figure 2.2. Schematic representation of size-dependent amyloid fibril growth. Monomers in solution (green rectangles) diffuse into the surface of fibrils where they get reversibly adsorbed (yellow rectangles) and rapidly migrate to the fibril ends. Accumulation of adsorbed molecules at the active sites (red arrows) takes place until reaching the minimal number of monomers required for integration and subsequent regeneration of the fibril end. For β_2m, this number was estimated from pre-steady-state kinetic measurements to be ~10 monomers[100]. Supported on the non-active site binding mechanism[100,101], we further assume that lateral migration of molecules is not rate-limiting so that the influx of monomers directly increases with fibril length as the attachment probability also increases. Accordingly, regeneration events take place faster in (A) long fibrils than in (B) short fibrils – time-lapse is represented by the blackened clock symbols on the right side. Helical representation of amyloid fibrils adapted from[103].

2.2.4 Mathematical model

Given the kinetics of nucleation and growth just defined, it is important to know how much and how fast amyloid fibrillization will take place. We will start by defining the mass of polypeptide aggregates, m, in the generic case of fibril formation using a mass of seeds m_0 and taking into account the fibril size distribution:

$$m = m_0 + \Delta m = a_\rho \sum_{i=0}^{\infty} n_i L_i \qquad (2.10)$$

In the second equality, a_ρ is the cross-section area multiplied by the mass density of amyloid fibrils, and is assumed to not change significantly with time; n_i is the number of fibrils with characteristic size L_i. It follows that the rate at which m increases with time is given by

$$\frac{\mathrm{d}m}{\mathrm{d}t} = a_\rho \sum_{i=0}^{\infty} \frac{\mathrm{d}n_i}{\mathrm{d}t} L_i + a_\rho \sum_{i=0}^{\infty} \frac{\mathrm{d}L_i}{\mathrm{d}t} n_i \qquad (2.11)$$

The first sum term accounts essentially for the formation of new amyloid nuclei containing n^* molecules (eq. 2.7), which occurs at a nucleation frequency J (eq. 2.8). The second sum term refers to the size-dependent elongation of fibrils (eq. 2.9). Fibril breakage, although important for the final size distribution and for pathogenic proliferation[104,105], is not expected by the CLM to influence the overall kinetics of m increase: on the one hand, fragmentation increases the number of fibrils but not their total mass; so, the first sum term of eq. 2.11 remains constant. On the other hand, provided that the initial mass is the same, a large number of shorter and slow growing fibrils are expected to have the same total mass increase per time unit as fewer and fast growing fibrils. This is due to the condition of size-dependent growth resulting from eq. 2.9, which implies that the same mass ofdifferent sized seeds will induce unvarying fibrillization kinetics. Experiments of amyloid fibril formation of β_2m using small amounts of seeds apparently confirm that short, ultrasonicated fibrils induced "essentially the same kinetics" as less fragmented ones[106]. Nevertheless, as higher α_0 fractions are used, fragmented seeds seem to promote the initial fibrillization rates v_0[68,100]; more surprisingly, when ultrasounds or mechanical agitation are used during incubation, the overall conversion into amyloid fibrils (expressed in terms of final fluorescence intensity) is markedly affected by the type of agitation[107,108]. From these results, we suggest the existence of a thermodynamic rather than kinetic effect associated to fibril breakage. The effective concentration of protein in solution is likely to be influenced by the total number of monomers remaining captive on the fibril ends (Figure 2.2). The bigger is fibril fragmentation, the larger is the number of fibril ends, and more affected is the effective concentration in equilibrium. Although important to explain unexpected behaviors, we will not be concerned with changing thermodynamics due to fibril breakage and/or volume-excluding effects at this stage of the CLM development. Equation 2.11 is therefore re-written as a function of the kinetics of nucleation and growth:

$$\frac{\mathrm{d}m}{\mathrm{d}t} = A\sigma^2 Vn^* \frac{M_\mathrm{w}}{N_\mathrm{A}} + k_g \sigma \sum_{i=0}^{\infty} a_\rho n_i L_i \qquad (2.12)$$

with M_w being the molecular weight of the amyloid protein and N_A the Avogadro constant. After recognizing that the sum term corresponds to the total mass m (eq. 2.10), and expressing Δm and σ as a function of the extent of amyloid fibril formation α (eqs. 2.1 and 2.6) one gets

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t} = k_\mathrm{a}\left(1-\alpha\right)\left[k_\mathrm{b}\left(1-\alpha\right)+\alpha+\alpha_0\right] \qquad (2.13)$$

where the σ_0-dependent constant k_a gives the magnitude of the growth step:

$$k_\mathrm{a} = k_g \sigma_0 \qquad (2.14)$$

and k_b gives the relative nucleation-to-growth magnitude:

$$k_\mathrm{b} = \frac{A\bar{n}^* M_\mathrm{w}}{N_\mathrm{A} C^* k_g} \qquad (2.15)$$

An average value $\bar{n}^*$ is considered for the number of molecules constituting the amyloid nuclei. Although n^* increases as $\Delta\mu$ (and σ) decreases (eq. 2.7), the average size of the critical nuclei is assumed for now to be dictated by the initial supersaturation σ_0. In eq. 2.13, α_0 corresponds to the mass fraction of seeds ($m_0/\Delta m_T$) and not to the initial amyloid conversion, which is 0 for both seeded and unseeded reactions $\alpha(0)=0$. The analytical solution of the ordinary differential equation 2.13 subject to this initial condition is

$$\alpha = 1 - \frac{1}{\frac{k_\mathrm{b}+\alpha_0}{1+\alpha_0}\left[\exp\left[k_\mathrm{a} t\left(1+\alpha_0\right)\right]-1\right]+1} \qquad (2.16)$$

which in the case of unseeded reactions simplifies to

$$\alpha = 1 - \frac{1}{k_\mathrm{b}\left[\exp\left(k_\mathrm{a} t\right)-1\right]+1} \qquad (2.17)$$

This equation is a major result of our study. Some of the practical and fundamental applications arising from it will now be discussed using experimental amyloid fibrillization kinetics obtained from the literature. Equation 2.17 will be fitted to the measured variations of α with time in order to determine the kinetic constants k_a and k_b. In so doing, the growth and nucleation steps will be quantitatively characterized, which is useful to design and predict different amyloidogenic conditions. When the rates of amyloid fibril formation are available for different initial protein concentrations (and for different σ_0), further

fundamental insight about amyloid fibrillogenesis is gained by disclosing the meaning of the kinetic constants at the molecular scale. Our future work will address the cases of seeded fibrillization kinetics (eq 2.16) in more detail.

2.2.5 Comparative study: Numerical fit of two-parameter models to experimental data on amyloid fibrillization kinetics

Unlike what happens in protein aggregation kinetics with perfect sigmoid shape, the hyperbolic type of kinetics is not rate-limited by the nucleation step. Therefore, in order to comprehensively compare the two-parameter models listed in Table 2.1, hyperbolic protein aggregation curves measured by Hurshman et al.[70] will be used. Ferrone's equation is not included in Table 2.1 since only a fraction of data points can be used during numerical fit of this model [55,74].

Table 2.1. Two-parameter models under comparison. Amyloid fibrillization kinetics are given by the fractional conversion of monomers into fibrils expressed computed as a function of time.

Model	Model Equation	Comments
CLM	$\alpha = 1 - \frac{1}{k_b\left[\exp(k_a t) - 1\right] + 1}$	- Present work - Parameters: k_a and k_b
Finke-Watzky[51,75]	$\frac{C_0 - C}{C_0} = 1 - \frac{1 + \frac{k_1}{k_2 C_0}}{1 + \frac{k_1}{k_2 C_0}\exp\left\{\left(1 + \frac{k_1}{k_2 C_0}\right) k_2 C_0 t\right\}}$	- C_0 and C are the initial and instantaneous protein concentration, so that $\frac{C_0 - C}{C_0} \approx \alpha$ - Parameters: k_1 and k_2
Saitô's [32]	$\alpha = \rho\frac{\exp\left[(1+\rho)kt\right] - 1}{\rho\exp\left[(1+\rho)kt\right] + 1}$	- Parameters: ρ and k
Fernández et al.[76]	$\alpha = \frac{1 - \exp(-k_{app} t)}{1 + a\exp(-k_{app} t)}$	- specifically derived for α-synuclein-polyamine complexes - Parameters: k_{app} and a

Non-linear curve fit was performed using Microcal Origin 6.0 software package based on the Levenberg-Marquardt algorithm of χ^2 minimization. The fitting results are given in Table 2.2 in terms of standard error (SE) of fitted parameters, minimized value of χ^2 and goodness of fit (R^2).

Table 2.2. Fitting the two-parameter models listed in Table 2.1 to monomeric transthyretin (M-TTR) aggregation kinetics measured by Hurshman et al.[70], digitized by us, and normalized by the fluorescence endpoint signals.

CLM						
mg/mL of M-TTR	$\chi^2\times10^4$	R^2	$k_a\times10^3$	SE (k_a)$\times10^4$	k_b	SE (k_b)
0.05	12	0.981	1.1	1.8	0.66	0.16
0.10	2.3	0.997	1.6	1.2	1.28	0.13
0.15	3.0	0.996	1.6	1.6	1.92	0.26
0.20	0.5	0.999	1.5	0.8	2.87	0.19
0.25	2.3	0.996	3.0	2.9	1.73	0.24
0.30	1.2	0.997	2.6	2.6	3.19	0.41
0.40	2.3	0.994	3.4	5.0	3.58	0.68
Finke-Watzky (F–W) Model						
mg/mL of M-TTR	$\chi^2\times10^4$	R^2	$k_1\times10^3$	SE (k_1)$\times10^4$	$k_2C_0\times10^3$	SE (k_2)$\times10^4$
0.05	12	0.981	0.73	0.7	0.38	2.3
0.10	2.3	0.997	2.0	0.7	-0.47	1.7
0.15	2.9	0.996	3.0	1.1	-1.5	2.6
0.20	0.5	0.999	4.4	0.7	-2.9	1.4
0.25	2.3	0.996	5.2	2.2	-2.1	4.9
0.30	1.1	0.997	8.1	2.5	-5.4	4.9
0.40	2.3	0.994	12	6.2	-9.0	11
Saitô's Model						
mg/mL of M-TTR	$\chi^2\times10^3$	R^2	ρ	SE (ρ)	$k\times10^5$	SE (k)$\times10^4$
0.05	1.2	0.981	1.9	1	38	2.4
0.10	0.39	0.995	37.6	181	5.0	2.3
0.15	1.0	0.987	57.6	736	4.0	5.5
0.20	1.3	0.981	141.6	5487	2.0	9.0
0.25	0.68	0.989	189.3	7177	2.0	8.6
0.30	1.8	0.959	265.7	23170	2.0	19
0.40	2.6	0.927	318.8	43030	2.0	33
Fernández et al. Model						

mg/mL of M-TTR	$\chi^2\times10^3$	R^2	$k_{app}\times10^3$	SE (k_{app})$\times10^4$	a	SE (a)
0.05	1.2	0.981	1.1	1.8	0.52	0.37
0.10	0.47	0.994	2.0	1.8	0.09	0.15
0.15	1.1	0.986	2.6	4.0	0.06	0.24
0.20	1.4	0.979	3.4	6.2	0.04	0.29
0.25	0.72	0.988	4.4	5.6	0.02	0.21
0.30	1.9	0.958	5.7	11	0.02	0.33
0.40	2.6	0.927	8.0	20	0.01	0.42

For the sake of numerical comparison, unconstrained fit was performed for all models under study. In the case of the F–W model, this procedure resulted in negative growth rate constants (k_2) which are physically unrealistic unless fibril dissolution is admitted. The results of Table 2.2 show that the best fitting results are obtained by the CLM and by the particular solution of the F–W model having no physical meaning. This is also illustrated in Figure 2.3, where lower values of χ^2 and higher values of R^2 are associated to the CLM and to the F–W model. In addition, the parameters of these two models are associated to lower standard errors (Table 2.1) meaning that the quality of fitting is more sensitive to parameter variations.

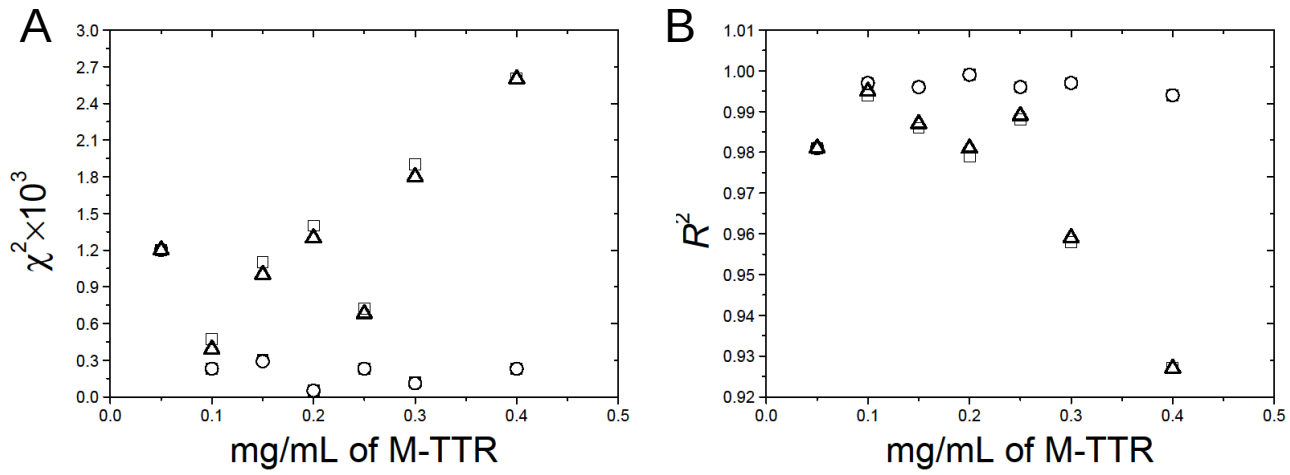


Figure 2.3. Quality of fitting expressed in terms of the values of (A) χ2 and (B) R2 for the different conditions listed in Table 2.2. Comparison between the CLM (○), F–W model (■), Saitô's model (Δ) and Fernandez et al. model (□). Results obtained by the CLM and F–W models are superimposed.

Since the equations of Saitô, Fernández et al. and F–W are interconvertible by algebraic manipulation, the following relationships between parameters can be established:

$$\rho = a^{-1} = \frac{k_1}{k_2 C_0} \qquad (2.18)$$

$$k = k_{app} = k_2 C_0 \qquad (2.19)$$

As a consequence, it is not surprising that the quality of fitting results of Saitô´s model and Fernández et al. model are very approximate – Figure 2.3 and Table 2.2. In the case of the F–W model, the algorithm of χ^2 minimization spontaneously converged to a better numerical solution (although lacking physical meaning) for all but one of the concentrations under study. Equally good agreement with the experimental data could be obtained by Saitô´s model and Fernández et al. model by adopting physically unrealistic, negative values for each parameter. Interestingly, the best numerical fit of the three models corresponds in terms of χ^2 and R^2values to the result obtained using the CLM. These results show that the solid biophysical basis of the CLM is also supported by better curve-fitting results.

2.3 Discussion

The CLM will be tested against experimental data published by Xue et al.[68] and Hurshman et al.[70] that comprehensively describe unseeded sigmoidal and hyperbolic kinetics. Although different numerical methods can be applied to determine the model parameters ka and kb, we propose a simplified strategy that uses the time required to reach 50% of the maximal amplitude ($t50$) and the aggregation rate at that instant ($v50$). The theoretical value of the product $k_a t50$ is obtained by setting $\alpha = 0.5$ in eq. 2.17:

$$k_a t_{50} = \ln\left(\frac{1}{k_b} + 1\right) \qquad (2.20)$$

In the same way, v_{50} corresponds to the derivative given by eq. 2.13 for $\alpha = 0.5$ and $\alpha_0 = 0$, with the product $v_{50}t_{50}$ being also a sole function of k_b:

$$v_{50}t_{50} = \frac{1}{4}(k_b + 1)\ln\left(\frac{1}{k_b} + 1\right) \qquad (2.21)$$

As shown in Figure 2.4A, v_{50} is obtained from the slope of the experimental $\alpha(t)$ curve at the instant t_{50}. In a first step, the value of k_b is determined from solving eq. 2.21 using the measured values of t_{50} and v_{50}. The value of k_a is then determined from solving eq. 2.20 using t_{50} and the k_b constant just estimated. This two-step method for determining the model parameters is graphically illustrated in Figure 2.4B.

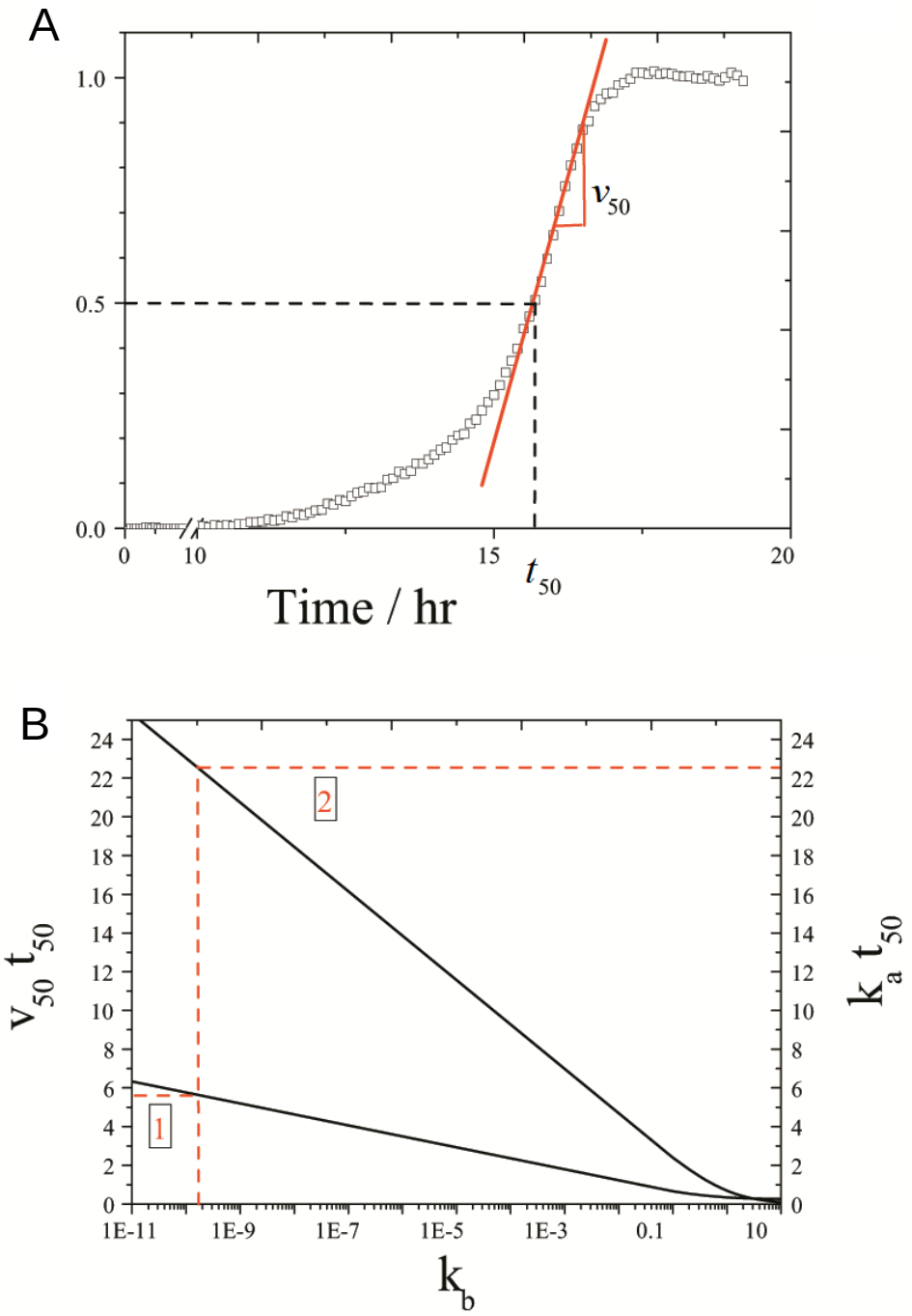


Figure 2.4. Simplified method for determining the kinetic constants k_a and k_b from unseeded amyloid aggregation kinetics. (A) Illustrative example for determining the time required to reach 50 % completion t_{50} (dashed lines) and the aggregation rate at the that instant v_{50} (solid lines) using the time dependent, normalized fluorescence signal measured by Xue et al.[68] for 61 μM of β_2m; t_{50}=15.6 hr and v_{50}=0.36 hr^{-1}. (B) The products $v_{50}t_{50}$ (left vertical axis) and k_at_{50} (right vertical axis) are computed as a function of k_b (log-scale) according to eqs. 2.21 and 2.20, respectively (solid lines). From the kinetic parameters determined in (A), k_b is firstly interpolated from the product $v_{50}t_{50}$, and then is used to interpolate the product k_at_{50} from which k_a is estimated (dashed arrows); k_b=1.6×10^{-10} and k_a=1.44 hr^{-1}.

The mechanistic similarities suggested by Fändrich[109] when analyzing 298 amyloid aggregation kinetics of different polypeptides find here their underlying principle. The author noticed that the product of the apparent lag time t_1 and the maximum aggregation rate v_{max} did not change significantly from an average value of 4.5 with standard deviation of ±2.9 [80,109]. Although t_1 is lower than t_{50}, and v_{max} is greater or equal to v_{50}, the product $t_1 v_{max}$ is expected to be close to $v_{50} t_{50}$. The representation of eq. 2.21 in Figure 2.4B shows that a relatively narrow range of $v_{50} t_{50}$ values, say between 2 and 6, corresponds to k_b values that vary many orders of magnitude, from ~10^{-11} to ~10^{-4}. That is to say, according to the CLM, very different nucleation-to-growth ratios exhibited by amyloid polypeptides have associated nearly invariable $v_{50} t_{50}$ and $t_1 v_{max}$ products. The bad news are that even small experimental errors during aggregation kinetics measurement can lead to great uncertainty on the estimated value of k_b. The values of k_b associated to the$t_1 v_{max}$ products identified by Fändrich[109] are much lower than 1, so, in eq. 2.21, $k_b + 1 \cong 1$. By combining eq. 2.20 and the simplified version of eq. 2.21 one finds that $k_a \cong 4 v_{50}$ for the large number of amyloid aggregation kinetics analyzed. The simplified method illustrated in Figure 2.4 was applied to more than 200 fibrillization progress curves measured by Xue et al.[68] at concentrations of β_2m monomer ranging from 8 to 244 μM. The selection of fitting results given in Figure 2.5A illustrates the good agreement obtained throughout the different polymerization stages and for all concentrations.

This is especially relevant if one recalls that half-life coordinates (t_{50} and v_{50}) were the only information used to compute the theoretical curves in their full extent. According to eq. 2.14, the values of k_a should increase linearly with the initial supersaturation level σ_0. Mean k_a values of replicate experiments are represented in Figure 2.5B as a function of the σ_0 values calculated using eq. 2.5 and a C^* estimation of 0.67 μM for β_2m under acidic conditions[110].

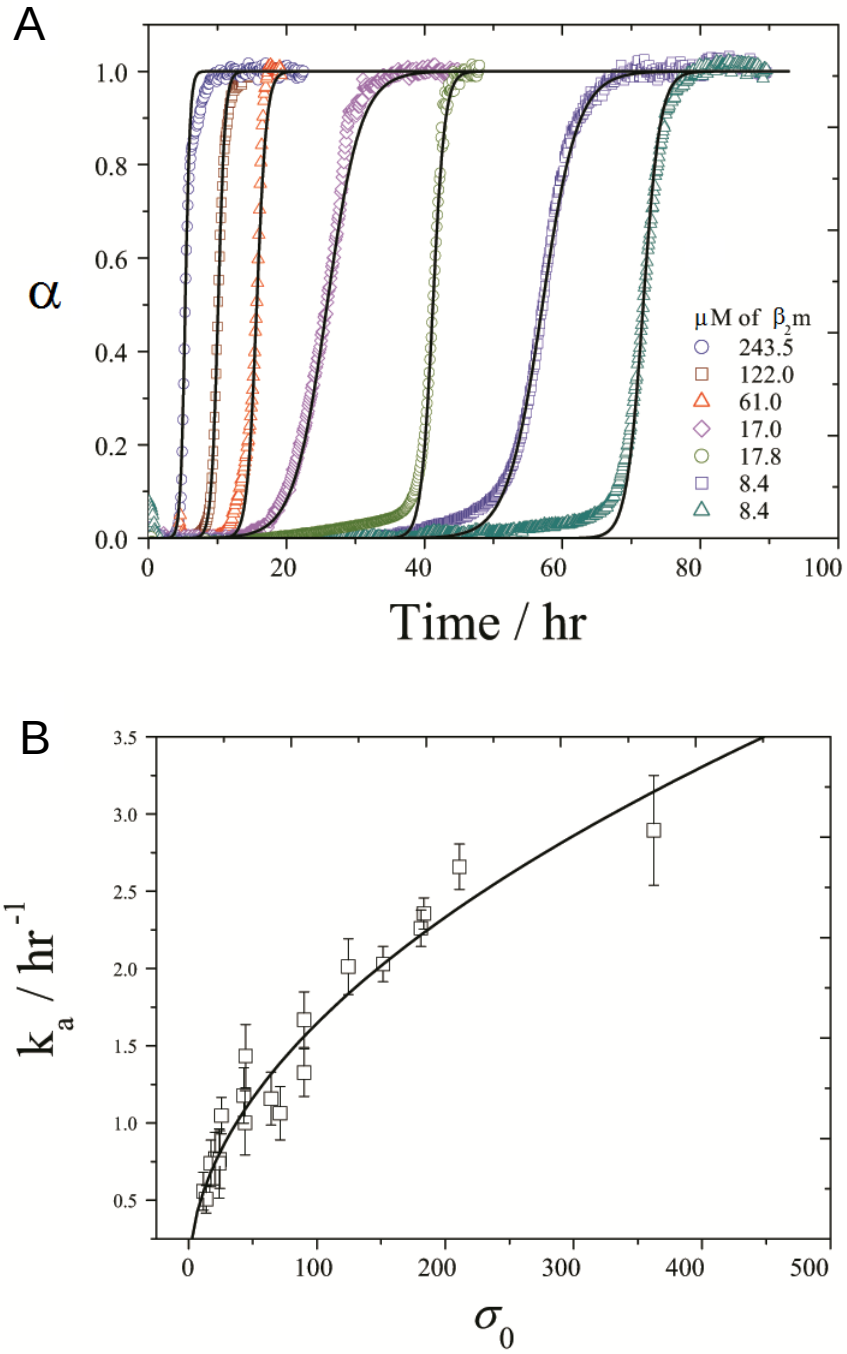


Figure 2.5. Fitting the CLM to sigmoidal aggregation kinetics measured by Xue et al.[68] at different β_2m concentrations. (A) Theoretical kinetic curves (lines) are calculated using eq. 2.17 and the k_a and k_b constants obtained by the simplified method described in Figure 2.4. The selection of results covers very fast to very slow β_2m fibrillization kinetics. (B) Influence of the initial supersaturation level σ_0 on the mean k_a values (symbols) determined for different protein concentrations within a 95 % confidence interval (error bars). Solid line represents a power law equation with exponent 0.503.

The fact that the $k_a(\sigma_0)$ relationship is closer to a power law function than to a linear function is not discouraging. On the contrary, it reflects the increasing diffusional limitations as the C_0 increases from 8 to 244 μM. More concentrated protein solutions result in a higher quantity of suspended fibrils, which in turn decreases the mass transfer rates and does not let k_a to increase with σ_0 as fast as in diluted solutions. Concerning the determination of the constant k_b, a great variability of results was obtained for each β_2m concentration that was analyzed. The 95% confidence intervals for the mean k_b included the value 0 in all cases except for C_0=43.7μM. The difficulty in finding statistically valid k_b estimations confirms our concerns about the high sensitivity of this parameter, at least in the region of Figure 2.4B corresponding to high values of $v_{50}t_{50}$. The same uncertainty does not let us to conclude about the meaning of some misfit occurring when α starts to increase using less concentrated β2m solutions – Figure 2.5A. These discrepancies, which were not reproducibly observed, may represent different nucleation rates than those considered by the CLM or, more simply, may result from experimental baseline variations that are more likely to occur during slower fibrillization kinetics.

So far we have been analyzing sigmoidal aggregation kinetics showing an initial lag phase that precedes the phase of rapid growth. As the nucleation step becomes not so rate-limiting, i.e., as k_b increases and $v_{50}t_{50}$ decreases, the initial lag phase becomes progressively shorter until no induction time is discernible. At the same time, the shape of the amyloid fibrillization kinetics is expected to gradually change from sigmoidal to hyperbolic. This is understandable from the value of amyloid fibril conversion at the inflection point α_i, which is calculated from the condition $d^2\alpha/dt^2 = 0$ applied to eq. 2.13 for unseeded reactions ($\alpha_0 = 0$):

$$\alpha_i = \frac{1-2k_b}{2(1-k_b)} \qquad (2.22)$$

For low values of the nucleation-to-growth ratio, the progress of amyloid fibrillization follows a sigmoidal shape with an inflection point at $\alpha_i = 0.5$. As k_b increases, the sigmoid curves have their inflection at lower values of α until a stage where no inflection point is visible at all. This happens for $k_b > 0.5$ and corresponds to the hyperbolic aggregation kinetics.

A good agreement between experimental and theoretical results is again observed in the case of hyperbolic aggregation kinetics of monomeric transthyretin (M-TTR) –

Figure 2.6A. First estimations of k_a and k_b using the values of t_{50} and v_{50} were now fine-tuned through numerical minimization of the squared error. The parameters thus obtained are in close agreement with those obtained using Levenberg-Marquardt algorithm of χ^2 minimization – Table 2.2. Figure 2.6B shows that the obtained values of k_a and k_b seem to linearly increase with the initial supersaturation level $\sigma_0$calculated using eq. 2.5; a value of 0.8 μg/mL was used for the M-TTR solubility C^*, which corresponds to the concentration of monomers in solution after 1 week of reaction time[70]. While the linear $k_a(\sigma_0)$ relationship confirms the definition of the growth rate constant given by eq. 2.14, the $k_b(\sigma_0)$ data suggests that, in eq. 2.15, the average number of monomers constituting the critical nucleus $\bar{n}^*$ increases with the initial supersaturation.

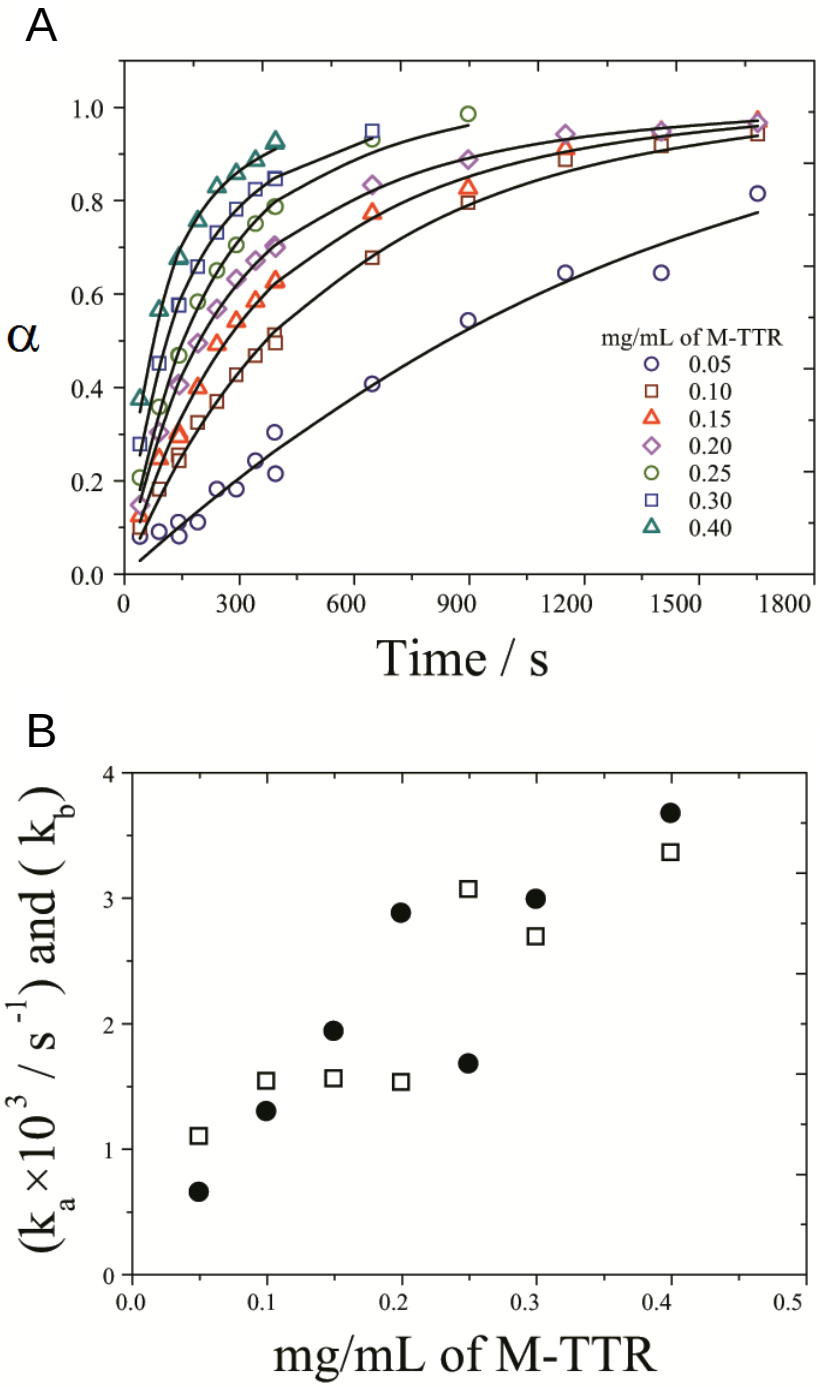


Figure 2.6. Fitting the CLM to hyperbolic aggregation kinetics measured by Hurshman et al.[70] at different monomeric transthyretin (M-TTR) concentrations. (A) Symbols represent continuous fluorescence data that was digitized by us into periodic data and normalized by the endpoint signal. Theoretical kinetic curves (lines) are calculated using eq. 2.17 and the k_a and k_b constants obtained by minimization of the squared error. (B) Influence of the initial supersaturation level σ_0 on the estimated k_a (open symbols) and k_b constants (full symbols).

This result is apparently in contradiction with the Classical Nucleation Theory which expects smaller critical nuclei as the thermodynamic driven force increases (eq. 2.7).

Additional statistically-reliable nucleation data is however needed before taking more definite conclusions about the definition of n^* for amyloidogenesis. M-TTR andβ_2m are amyloidogenic proteins that have rather distinct fibrillization kinetics, namely in reaction speed and shape of the progress curve. The ability of CLM to describe such different aggregation mechanisms is very encouraging. Another feature clearly distinguishing the two behaviors is the role of seeding on the acceleration of amyloid aggregation. While the addition ofpreformed aggregates of β_2m has the typical effect of precluding the initial lag phase[68], in the case of M-TTR the aggregation kinetics is not accelerated by seeding[70]. The CLM again provides the rationale for these observations based on the relative importance of the nucleation step relatively to growth: the very low values of $k_b$estimated for β_2m mean that amyloid nucleation is the rate limiting step. Therefore, the initial fibrillization rate v_0 predicted by eq. 2.13 for $\alpha = 0$ is also very low in the absence of seeding. Even small amounts of seeds may represent mass fractions α_0 much higher than k_b, which will lead to a correspondingly high increase on the value of v_0 and to a change of fibrillization kinetics from sigmoidal to hyperbolic. On the contrary, the high values of k_b obtained for M-TTR represent fast fibrillization rates right from the beginning of the reaction. This occurs independently of the values of α_0 that are typically well below 1. Drastic differences in the value of k_b may also configure different mechanisms of M-TTR and β_2m amyloid fibril formation: amorphous M-TTR aggregates identified by microscopic analysis[70] are likely act as nucleation promoters[111]. Instead, β_2m fibrils had the long-straight morphology typical of low pH, low ionic-strength conditions[68,112], suggesting that elongation is the major route for fibril formation.

At this point the CLM is confirmed as a robust two-parameter model that is able to describe very distinct protein aggregation kinetics and whose derivation is theoretically consistent and not liable to oversimplified hypothesis. As shown in Table 2.2 and Figure 2.3, the CLM also obtained the best fitting results among the two-parameter models used to describe protein aggregation data that is not rate-limited by any of the steps. In a subsequent stage, additional mechanistic insight into the nucleation and growth steps should be gained by disclosing the meaning of the kinetic constants k_a and k_b. Our preliminary study on how these constants are affected by the initial supersaturation level seems to indicate that modifications to the Classic Nucleation Theory are needed in order to determine the critical size of the amyloid nucleus. Concerning the growth step, the condition of size-dependent growth inherent to eq.2.9 also deserves further discussion.

This condition is not easily tested by direct observation techniques, which are known to have experimental limitations such as constrained assembly on the mica/glass surface and growth rate dispersion[39,91–95]. An indirect, yet distinctive consequence of size-dependent growth is that fibrillization rates are affected by the total mass of fibrils present in solution but not by their number. Increased seeding efficiency obtained with fragmented (smaller) seed fibrils has however been observed for β_2m [68,100] and for yeast prions[113], which is in apparent contradiction with the size-dependent condition. Fibril breakage induced by different agitation intensities have also been reported to affect the end-point fluorescence signal[107,108], which is an indication of the total amount of fibrils produced and, indirectly, of the protein solubility (eq. 2.6). Both type of evidences are not expected by the CLM and are presumably related with variations on the effective protein concentration due to volume excluding effects in fibril-crowded solutions. Known as an important variable in cellular environment, macromolecular crowding and its effects are generally associated to high concentrations of spectator molecules[114]. We hypothesize that the thermodynamic concentration of protein may also be affected *in vitro* by fibril breakage as an additional source of monomer-retaining sites. Besides providing a possible explanation for masked fibrillization kinetics, changing thermodynamic conditions may also be linked to a series of unusual behaviors reported in literature such as the effects of reaction volume[115], hydrophobic surface area[116] and surfactants[108] on amyloid fibril formation. In all cases, not only the reaction rate is affected but also the reaction extent, which again supports the thermodynamic hypothesis of volume excluding effects taking place at the fibril-solution interface. Whether the effect of fibril breakage is thermodynamic or kinetic is important to evaluate the condition of size-dependent growth and to establish the fundamental meaning of constant k_a: (i) is it related to probabilistic interfacial collisions and the mean life time of monomers at the surface of amyloid fibrils; (ii) or should it be treated as a blind kinetic constant such as those commonly found in protein aggregation models that consider reaction-like rate equations? Either way, sufficient experimental data is available confirming a first-order dependency of the elongation rate on the fundamental variables supersaturation level and fibril mass/size, which in turn leads to a first-order dependency of k_a on the initial supersaturation level σ_0.

The CLM opens new perspectives on the characterization and understanding of amyloid fibril kinetics. Development of new targets for amyloid disease therapeutics needs this model working together with other structural and molecular studies[102,117] and

complemented with information on the composition of toxic nonfibrillar oligomers[86,93]. CLM allows to discern between the following mechanisms of inhibition: (i) protein solubility increase evidenced by higher values of C^* in the presence of the inhibitor thus lowering the thermodynamic driving force for fibrillization – eq. 2.5; (ii) nucleation rates decrease evidenced by lower nucleation-to-growth ratios (k_b) in the presence of the inhibitor; (iii) blocking of fibril elongation evidenced by lower kinetic constant for growth (k_a) in the presence of inhibitor. Mechanisms that require structural interaction between ligand and polypeptide, such as nucleation and/or growth inhibition offer more guarantees to work both *in vitro* and *in vivo*. By characterizing the relative contribution of variables, the CLM is expected to be useful for scientific systematization during the development of new therapeutic strategies for amyloid diseases.

In conclusion, we propose a quantitative Crystallization-Like Model that is able to describe the sigmoid and hyperbolic amyloid fibrillization kinetics commonly reported in literature. A simplified method is presented to determine the model parameters using Figure 2.4B and the measured values of the time required to reach 50% completion (t_{50}) and the aggregation rate at the that instant (v_{50}). Further insights are provided into the nature of critically sized amyloid nucleus, the influence of diffusive limitations and fibril breakage during growth, and the role of seeding as amyloid aggregation catalyst. It is notable that the solid biophysical basis of the CLM does not compromise its final simplicity. While integrating the combined influence of seeding, nucleation, growth and fibril breakage, eq. 2.16 is seen as a good approximation to the generalized underlying principle governing the kinetics of amyloid fibrillization. We expect that the CLM may be routinely used as a tool to characterize quantitatively and qualitatively new targets for amyloid disease therapeutics.

Acknowledgments

We thank Wei-Feng Xue, Sheena E. Radford and their research team for allowing us to reproduce published experimental data in Figures 2.1, 2.4A and 4A. We thank Jeffery W. Kelly and his research team for allowing us to reproduce published experimental data in Figures 2.1 and 2.6A.

Chapter 3

How to screen amyloid off-pathway modulators

What can the kinetics of amyloid fibril formation tell about off-pathway aggregation?

Rosa Crespo[1], Eva Villar-Alvarez[2], Pablo Taboada[2], Fernando A. Rocha[1], Ana M. Damas[3] and Pedro M. Martins[1,3]

[1]From LEPABE, Laboratório de Engenharia de Processos, Ambiente, Biotecnologia e Energia, Departamento de Engenharia Química, Faculdade de Engenharia da Universidade do Porto. Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

[2]Departamento de Física de la Materia Condensada, Facultad de Física, Universidad de Santiago de Compostela, Spain

[3]ICBAS – Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto, Portugal

To whom correspondence should be addressed: Pedro M. Martins, ICBAS – Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Rua de Jorge Viterbo Ferreira n.º 228, 4050-313 Porto, Portugal, Tel.: (+351) 220428137; E-Mail: pmmartins@icbas.up.pt

Keywords: Amyloid, protein aggregation, biophysics, mathematical modeling, kinetics, drug screening

Summary

Some of the most prevalent neurodegenerative diseases are characterized by the accumulation of amyloid fibrils in organs and tissues. While the pathogenic role of these fibrils is not totally established, increasing evidences suggest off-pathway aggregation (OPA) as a source of toxic/detoxicating deposits that still remains to be targeted. The present work is a step ahead towards the development of off-pathway modulators using the same amyloid-specific dyesas those conventionally employed to screen amyloid inhibitors. We identified a series ofkinetic signatures revealing the quantitative importance of OPA relatively to amyloid fibrillization; these include non-linear semi-log plots of amyloid progress curves, highly variable endpoint signals and half-life coordinates weakly influenced by concentration. Molecules that attenuate/intensify the magnitude of these signals are considered promising off-pathway inhibitors/promoters. An illustrative example shows that amyloid deposits of lysozyme are only the tip of an iceberg hiding a crowd of insoluble aggregates. Thoroughly validated using advanced microscopy techniques and complementary measurements of Dynamic Light Scattering (DLS), Circular Dichroism (CD) and soluble protein depletion, the new analytical tools are compatible with the high-throughput methods currently employed in drug discovery.

3.1 Introduction

Nucleation and growth mechanisms are ubiquitous in nature, from atmospheric aerosol formation[118] to microtubule polymerization in eukaryotic cells[119] and biomineralization[120]. The present work focuses amyloid fibril formation as a particular case of such phase transition phenomena. Amyloid fibrils are filamentous assemblies of proteins in which the polypeptide backbone is arranged in a characteristic cross-β-sheet structure running perpendicular to the long axis of the fibrils. Despite the conflicting evidence about its pathogenic role[121], the accumulation of amyloid deposits in a variety of organs and tissues is associated to the most prevalent neurodegenerative diseases and to amyloidosis[122]. Intermediate oligomeric species and off-pathway end-products may, however, be more dangerous for the development of amyloid diseases than mature fibrils themselves[1,121,123–125]. Conversely, the aggregation pathway might be redirected into insoluble oligomers less toxic than amyloid fibrils[126], given that both types of precipitates kinetically compete for soluble protein[13]. As non-fibrillar species are also hard to detect, the molecular mechanisms involved in their formation are particularly difficult to infer[127–129]. Their existence should, at any rate, be a disturbance to the otherwise predictable sigmoidal to hyperbolical kinetics of phase transition.

By lacking the cross-β-sheet structure, intermediate and off-pathway species cannot be identified using amyloid-specific dyes such asfluorescent thioflavin-T (ThT). If solely amyloid-like fibrils are produced, the development of the mass-of-aggregates signal will be equivalent to the depletion of protein monomers in solution[34]. Both macroscopic indicators should be the result of "direct" phase transition events of primary (concentration-driven) nucleation, secondary (autocatalytic) nucleation and fibril growth (or elongation). Other "indirect" events, including fibril fragmentation and association, do not involve the transition of molecules between phases and, therefore, do not affect the reciprocity between aggregation of fibrils and depletion of protein. Proposed in 2012, the "crystallization-like model" (CLM) describes how the molecular mechanisms of nucleation and growth affect the evolution of amyloid conversion (α) with time (t). The two-parameter CLM equation for unseeded reactions reads[6]:

$$\alpha = 1 - \frac{1}{k_b\left(e^{k_a t} - 1\right) + 1} \qquad (3.1)$$

where k_a, originally called the growth rate constant, is more broadly defined as the autocalytic rate constant since it can also include the contribution of secondary nucleation steps proportional to the amount of amyloid fibrils in solution. The parameter k_b gives the relative rates of primary nucleation steps over autocatalytic steps. Once demonstrated the ability of the CLM to describe traditional fibrillization kinetics[6] and discriminate between true and apparent amyloid inhibitors[130], new CLM-based tools are hereby proposed in order to identify off-pathway processes. A preliminary note is, however, required in order to define what are the regular kinetic behaviors expected in the absence of parallel aggregation.

The current understanding of the kinetics and thermodynamics of protein aggregation was comprehensively reviewed by Morris, Watzky and Finke[51], and more recently, by Gilliam and MacPhee[34]. The standard sigmoidal growth curve exhibiting an initial lag-phase followed by a period of rapid growth and a final plateau phase is, in general, satisfactorily fitted by the existing theoretical models, which gives rise to different explanations of the same result. There are, however, a number of other common behaviors that remain unexplained, even by the moresophisticated models[34]. Some of these inconsistencies are marked in red in Figure 3.1, where a graphical account of typical experimental results is also provided. The hyperbolic (concave) profiles exhibiting no inflection point or lag-phase in absence of seeding (Fig. 3.1a) are frequently reported in literature during the aggregation of e.g., serum albumin[131,132], transthyretin[70,133], β_2-microglobulin[134], the four repeat domain of Tau[135], apolipoprotein[136] and amyloid-β variants[137,138]. This kind of results is explained by the CLM as the result of predominant primary nucleation over the autocatalytic processes, and is also well fitted by the "Ockham's Razor" minimalistic model[51,139]. According to Oosawa-type models[140] such as those of Ferrone et al.[55] and Knowles et al.[56], the early stage increase of the mass of fibrils cannot occur any faster than by polynomial t^n or exponential $\exp(kt)$ lawsduring unseeded reactions[34,141]. Therefore, Oosawa-type models fail to explain the concave profiles of the form $1-\exp(-kt)$ shown in red in Figure 3.1A[34]. The CLM advantageously uses fundamental principles of phase transition to explain the nonlinear trends in Figure 3.1B. By expressing the driving force for protein aggregation as the thermodynamic supersaturation $\sigma = \left(C - C_a^*\right)/C_a^*$ (an approximation to the variation in chemical potential), the steady-state monomer concentration is expected to correspond to the amyloid solubility

$C_\infty = C_a^*$. At the same time that the CLM was being proposed, Yoshimura et al. urged the need to recognize amyloidogenicity as a property determined by the monomer concentration relative to solubility[3]; since then, a wealth of new evidences unanimously confirm supersaturation as a major driving force for protein aggregation[1,8–13]. Although this parallel with crystallization had been hinted before[40,70], the common practice in literature models is to assume the monomer concentration alone as the driving force for phase transition[51,55,56,139,140] meaning that, e.g., amyloid fibrils would continue to grow until the solution became completely depleted from soluble protein. In addition, the duration of the lag-phase would be a linear function of C_0 when represented in a log-logscale; this follows from model equations of variable complexity that, in general, can be simplified to a power law equation with a constant scaling exponent γ[34,142]. The same does not apply for the CLM, where the lag-times scale exponentially with supersaturation σ_0 rather than with the initial concentration C_0. The result is a "broken" curve such as those represented in Figure 3.1B, in which the scaling factorγ changes its value as C_0 decreases to values closer to C_a^*. Note that although the theoretical range of concentrations includes the solubility value, the formation of fibrils by primary nucleation is only expected to occur for C_0 values above a critical concentration higher than C_a^*. The existence of breaks in lag-time versus concentration plots is widely reported in literature as reviewed by Eden et al.[142]. The different profiles simulated in Figure 3.1B result from the influence of the initial supersaturation σ_0 on the primary nucleation rate constant k_b. Both k_a and k_b influence the time needed to reach 50% completion (t_{50}) as determined by the following equation[6]:

$$t_{50} = \frac{1}{k_a}\ln\left(\frac{1}{k_b}+1\right) \qquad (3.2)$$

While the parameter k_a is proportional to σ_0, the parameter k_b is proportional to the critical number of monomers constituting the amyloid nucleus n_c, whose variation with σ_0 is not so well established[6]. Determining the k_b vs. σ_0 relationship is particularly difficult owing to the amplified uncertainty associated to k_b estimates when sigmoidal aggregation curves are used[6]. Apparently in contradiction with the Classical Nucleation Theory, the value of k_b (and n_c) estimated from hyperbolic aggregation kinetics of transthyretin is proportional to σ_0 [6,70]. The inset in Figure 3.1B shows how different values of k_b/σ_0 influences the scaling exponent γ extracted for high monomer concentrations.

Sufficiently low k_b/σ_0 values are chosen in order to assure meaningful durations of the lag-phase. Although other scaling behaviors result from admitting different $k_b(\sigma_0)$ functions, the interval of γ valueslying between −2 and −1 is in good agreement with recent Monte Carlo simulations that extend the classical nucleation-elongation-fragmentation scheme to include a stochastic nucleation step[142]. Each of the closed-form solutions of Oosawa-type models predicts a different scaling exponent that remains approximately constant over the concentration range[34,143]; explaining the existence of broken curves was shown to require additional fundamental insights besides the simple interplay between the different mechanistic alternatives[142,144]. In Ockham's Razor-type models, sigmoidal aggregation curves exhibiting pronounced lag-phases correspond to the limit case of very low primary nucleation rates for which the model equations simplify to a logistic function[6]. Here again, the exponential factor γ is approximately constant (close to −1) over the concentration ranges that produce lag-phases. A striking result not covered by any of the scenarios in Figure 3.1B is the weak dependence of the lag-time on concentration with absolute values of γ well below 1 at high monomer concentrations[34,56,142,145]. In the case of lysozyme aggregation under harsh denaturating conditions the process becomes nearly concentration independent ($\gamma \simeq 0$) for protein concentrations significantly higher than the solubility[142]. The formation of intermediate and off-pathway species is a likely explanation of this and of other deviations from canonical kinetics as illustrated next for HEWL aggregation under conditions of low pH (1.6) and high temperature (60 ºC) known to produce other aggregates than amyloid fibrils[146–148].

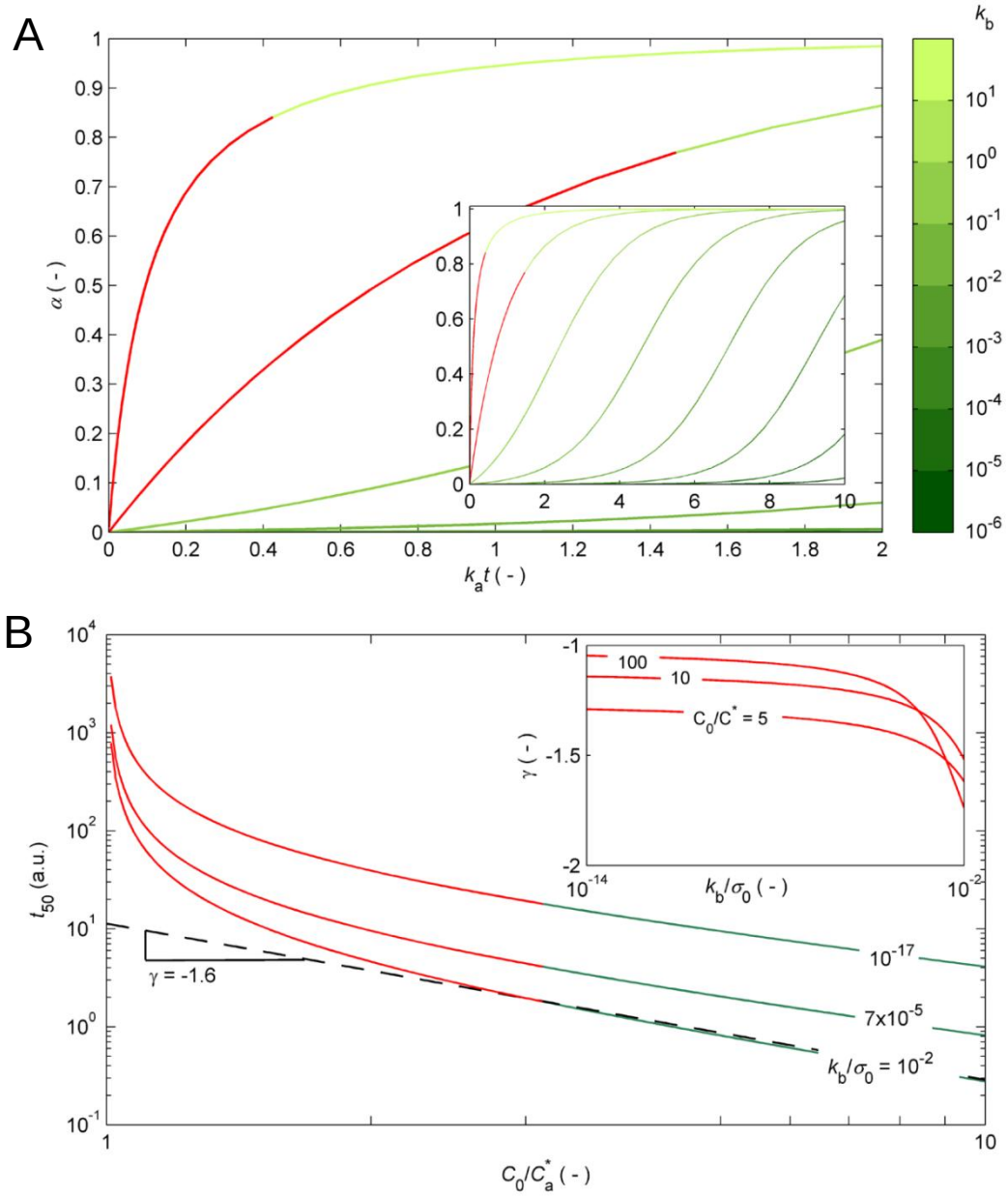


Figure 3.1. Canonical kinetic profiles expected by the CLM (green and red) but not expected by other theoretical models (red). (A) Hyperbolic to sigmoidal protein aggregation curves are obtained from Eq. 1 as the relative magnitude of primary nucleation decreases from k_b=10 to 10^{-6} (log-scale color bar). Inset: the $k_a t$ timescale is expanded to show complete sigmoidal growth curves. The hyperbolic profiles marked in red are not expected for unseeded reactions by Oosawa-type models[55,56,140]. (B) log-log representation of t_{50} (an indicator of the duration of the lag phase) as a function of the protein concentration computed using Eq. 2 for different values of k_b/σ_0. The slope of the dashed line corresponds to the exponential scaling factor γ (illustrative example for high protein concentrations and $k_b/\sigma_0=10^{-2}$). The broken curves shown in red are not expected by the different closed-form solutions of Oosawa-type and Ockham's Razor-type models. Inset: in the absence of off-pathway processes, the absolute value of γ is comprised between 1 and 2 according to the value of k_b.

3.2 Experimental Procedures

Chemicals:

HEWLwas obtained from Merk KGaA (Darmstadt, Germany). ThT was obtained from Sigma-Aldrich (St Louis, MO, USA) and used as received. Other chemicals were reagent grade and obtained from Merk KGaA (Darmstadt, Germany).

HEWL preparation:

HEWL powder was dissolved in 25 mM HCl pH 1.6 and dialyzed against 25 mM HCl, pH 1.6, using a cut-off 3500 Da membrane (Spectrum, Fisher Scientific). The concentration of the dialyzed protein solution was determined by absorbance measurements at 280 nm using an extinction coefficient of 37752 $M^{-1}cm^{-1}$. Protein stocks were stored at 4 ºC for no longer than one week.

ThT Fluorescence:

ThT fluorescence kinetic measurements were carried out at 60 ºC in 96-well plates (Thermo Scientific, microtiter) in a CHAMELEON ^{TM}V Microplate Reader (Hidex Co., Turku, Finland) at an excitation wavelength of 440 nm and an emission wavelength of 485 nm. ThT stock solution was prepared by dissolving the dried powder in 25 mM HCl, pH 1.6, and filtered through a sterile 0.45 μm pore size PES membrane filter (Jet Biofil). The concentration was determined by absorbance measurements at 411 nm using an extinction coefficient of 22000 $M^{-1}cm^{-1}$. Samples of 120 μl with a final ThT concentration of 2.8 mM and HEWL concentrations of 0.60, 0.93, 1.25, 1.39 and 1.76 mM were sealed with 100 μl of paraffin oil. Measurements were recorded every 1800 s, after sample homogenization by 300s shaking. Data was background-corrected for the ThT fluorescence of the respective solvent in the absence of protein.

Depletion of soluble HEWL:

Independent 1.76 mM HEWL samples were incubated at 60 ºC and periodically filtered through a sterile 0.22 μm pore size PES membrane filter (Jet Biofil). After 1 hour at room temperature, filtered samples were diluted 1:50 in 25 mM HCl pH 1.6 and analyzed spectrophotometrically at 280 nm using a 1 cm pathlength quartz cuvette (Hellma GmbH & Co.KG, Müllheim, Germany) and an extinction coefficient of 37752 $M^{-1}cm^{-1}$.

CD:

CD experiments were performed using a Jasco J-815 spectropolarimeter (Tokyo, Japan) equipped with a Peltier-controlled thermostated cell support. Independent 1.76 mM HEWL samples were incubated at 60 ºC for periods of 1 to 7 days. Samples,were diluted 1:400 in 25 mM HCl, pH 1.6. CD spectra were measured from 190 to 260 nm in a 0.1 cm pathlength quartz cuvette (Hellma Analytics). The final spectrum of all samples was an average of 16 independent scans recorded with 1 nm bandwidth, 2s digital integration time and a scanning speed of 50 nm/min.

DLS:

DLS measurements were performed at 25 ºC using an ALV/DLS/ SLS-5000F, SP-86 goniometer system (ALV-GmbH, Langen, Germany) equipped with a CW diode-pumpedNd:YAG solid-state Compass-DPSS laser with a symmetrizer (Coherent Inc., Santa Clara, CA). The laser operates at 532 nm with an output power of 400 mW. The intensity scale was calibrated against scattering from toluene. Independent 1.76 mM HEWL samples incubated at 60 ºC for periods of 1 to 7 days were analyzed at least three times. Measurements were made at a scattering angle 90º to the incident beam for 5-10 min.

AFM:

Independent 1.76 mM HEWL samples incubated at 60 ºC for 0, 1, 2, 3, 4 and 7 days were diluted 1:400 in 25 mM HCl, pH 1.6. Samples were spin-coated onto silicon wafers and dried in vacuum conditions for 4-5 hours. AFM images were recorded in non-contact mode using an AFM, JEOL instrument (JSPM 4210) equipped with a nitride cantilever NSC15 from MicroMasch, USA. Typical working frequency and spring constant: 325 kHz and 40 N/m, respectively. Topography images were recorded adapting the offset point according to the roughness of each sample.

TEM:

Independent 1.76 mM HEWL samples were incubated at 60 ºC for periods of 1 to 7 days. Sample solutions were applied onto a carbon–formvar coated 200–400 mesh spacing grids. After 1 minute, excess sample solution was removed by blotting with filter paper and stained with filtered aqueous solution of 2% uranyl acetate (for 45 seconds). Grids were

examined under a JEOL JEM 1400 TEM (Tokyo, Japan) operated at 80 kV. Images were digitally recorded using a Gatan SC 1000 ORIUS CCD camera (Warrendale, PA, USA).

Mathematical Model Derivation:

In the presence of off-pathway aggregation, the increase of total aggregates in solution is due to the formation of either amyloid fibrils or off-pathway aggregates:

$$\frac{dm}{dt} = \frac{dm_a}{dt} + \frac{dm_{off}}{dt} \qquad (3.3)$$

While the formerprocess is described by the two-parameter CLM[6],the latter is here characterized by a off-pathway nucleation step, which, similar to amyloid nucleation, is considered to be second-order in relation to supersaturation:

$$\frac{dm_{off}}{dt} = k_{off}^{*} V \sigma_{off}^{2} \qquad (3.4)$$

The rate constant k_{off}^{*} is the nucleation frequency expressed per volume of solution (V). Since the two types of aggregates in question have distinct structural organizations, the concentration of soluble protein equilibrating the solid phase is also different, i.e., supersaturation has to be differently defined as a function of the protein solubility C_i^{*} for amyloid fibrils (subscript $i = a$) and for off-pathway aggregates (subscript $i = off$):

$$\sigma_i = \frac{C - C_i^{*}}{C_i^{*}} \qquad (3.5)$$

As the formation of total aggregates and the depletion of soluble protein are complementary processes, m is proportional to the difference $(C_0 - C)$ and σ_i is alternatively expressed as

$$\sigma_i = \left(1 - \frac{m}{M_i}\right)\sigma_{0,i} \qquad (3.6)$$

where M_i is the total amount of species *i*(either *a* or *off*) that would be produced in the absence of the other species (either *off* or *a*), and $\sigma_{0,i}$ is the initial supersaturation *i* evaluated according to eq. 3.5–both M_i and $\sigma_{0,i}$ are proportional to the difference $\left(C_0 - C_i^{*}\right)$. Accordingly, eq. 3.4 is rewritten as

$$\frac{dm_{off}}{dt} = k_{off}^{*} V \sigma_{0,off}^{2} \left(1 - \frac{m}{M_{off}}\right)^{2} \qquad (3.7)$$

whereas the differential form of the CLM equation[6],

$$\frac{\mathrm{d}\alpha_a}{\mathrm{d}t} = k_a \frac{\sigma_a}{\sigma_{0,a}} \left(k_b \frac{\sigma_a}{\sigma_{0,a}} + \alpha_a \right) \qquad (3.8)$$

is rewritten as

$$\frac{\mathrm{d}\alpha_a}{\mathrm{d}t} = k_a \left(1 - \frac{m}{M_a} \right) \left(k_b \left(1 - \frac{m}{M_a} \right) + \alpha_a \right) \qquad (3.9)$$

where k_a is defined as the autocalytic rate constant and k_b gives the relative rates of primary nucleation steps over autocatalytic steps. The amyloid conversion is normalized as $\alpha_a = m_a / M_a$ thus implying that the final value of 1 is only reached when no off-pathway aggregates are formed. In all other occasions the final value of α_a represents the fraction of amyloid fibrils produced in relation to that expected in the absence of a competitive process. In our simulations amyloid fibrils are not supposed to dissolve once the protein concentration decreases below the solubility value C_a^*. After that limit, the amyloid supersaturation σ_a is considered 0. By adopting normalized units of time ($\theta = k_a t$) and of total amount of aggregates ($\beta = m / M_0$) on eq, 3.9 results:

$$\frac{\mathrm{d}\alpha_a}{\mathrm{d}\theta} = \left(1 - \beta \frac{M_0}{M_a} \right) \left(k_b \left(1 - \beta \frac{M_0}{M_a} \right) + \alpha_a \right) \qquad (3.10)$$

Finally, the same procedure is applied to eq 3.4, which is then replaced in eq. 3.3 to obtain

$$\frac{\mathrm{d}\beta}{\mathrm{d}\theta} = \frac{M_a}{M_0} \left(1 - \beta \frac{M_0}{M_a} \right) \left(k_b \left(1 - \beta \frac{M_0}{M_a} \right) + \alpha_a \right) + \\ + k_{off} \left(1 - \beta \frac{M_0}{M_{off}} \right)^2 \qquad (3.11)$$

where k_{off} is the normalized rate constant for off-pathway aggregation,

$$k_{off} = \frac{k_{off}^* V \sigma_{0,off}^2}{k_a M_0} \qquad (3.12)$$

and M_0 is the final amount of total aggregates. Being proportional to C_0 minus the final concentration of soluble protein (either C_a^* or C_{off}^*), M_0 also corresponds to either M_a or M_{off} depending on which species – amyloid fibrils or off-pathway aggregates – dictates the equilibrium.

The system of ordinary differential equations formed by eqs. 3.10 and 3.11 was solved using Matlab®subject to the initial conditions $\alpha_a(0)=0$ and $\beta(0)=0$,and according to the details given in each of the following numerical simulations.

Simulation 1–Effect of the off-pathway rate constant k_{off} on the aggregation kinetics:

We have solved eqs. 3.10 and 3.11 for different combinations of parameters k_b and k_{off}. The used value of M_0/M_a was estimated based on measurements of $C_a^*=1.50$ mM and $C_{off}^*=0.75$ mM obtained during protein depletion experiments carried out at $C_0=1.76$ mM:

$$\frac{M_0}{M_a}=\frac{C_0-C_{off}^*}{C_0-C_a^*} \qquad (3.13)$$

Since off-pathway aggregates have the lowest solubility, the final amount of total aggregates M_0 is calculated using C_{off}^* as the equilibrium protein concentration and $M_0/M_{off}=1$.

Simulation 2–Effect of the initial protein concentration C_0 on the aggregation kinetics:

Equations 3.10 and 3.11 were also solved taking into account C_0 -dependent variables M_0/M_a (eq. 13), k_b (proportional to $\sigma_{0,a}$) and k_{off}, whose definition can be rewritten as

$$k_{off}=\frac{k_{off}^*}{M_w\kappa}\left(\frac{\sigma_{0,off}}{\sigma_{0,a}}\right) \qquad (3.14)$$

obtained after replacing in eq. 3.12 (i) the direct dependence of k_a on $\sigma_{0,a}$,

$$k_a=\kappa\sigma_{0.a} \qquad (3.15)$$

and (ii) the definition of M_0,

$$M_0=\left(C_0-C_{off}^*\right)M_wV \qquad (3.16)$$

with M_w being the molecular weight of the amyloidogenic protein.From the obtained $\alpha_a(\theta)$ progress curves we computed the corresponding $\alpha(\theta)$ curves following the definitions given in Simulation 1. The theoretical half-life coordinates t_{50} and v_{50} were

respectively obtained from the normalized time required to reach 50% conversion θ_{50} and from slope $\mathrm{d}\alpha/\mathrm{d}\theta$ at the same instant:

$$t_{50} = \frac{\theta_{50}}{\kappa\sigma_{0,a}} \qquad (3.17)$$

$$v_{50} = \kappa\sigma_{0,a}\left.\frac{\mathrm{d}\alpha}{\mathrm{d}\theta}\right|_{50} \qquad (3.18)$$

As arbitrary units of time and aggregation rate are used, an arbitrary value of κ is also considered. The influence of C_0 on $\sigma_{0,a}$ and $\sigma_{0,off}$ in eqs.3.14, 3.17 and 3.18 is given in eq. 3.5. Different solutions were obtained for different values of C_0, $k_b/\sigma_{0,a}$ and $k_{off}\sigma_{0,a}/\sigma_{0,off}$.

3.3 Results

The amyloid fibrillization curves of HEWL suggest OPA:

The comparative analysis made in Introduction indicates that the CLM can be used as a touchstone to identify andcharacterize unconventional aggregation kinetics. This is now illustrated using *in vitro* experimental data measured by us during the aggregation of HEWL under conditions of low pH (1.6), high temperature (60 ºC) and high HEWL concentration (≥ 0.60 mM) required to produce amyloid fibrils without added denaturants or salts. Representing the results as the normalized intensity of ThT fluorescence with time (Fig. 3.2a) and by the concentration-dependent half-life coordinates (Figs. 3.2b and 3.2c) does not seem to indicate especially unusual kinetics, except the weak concentration dependence of t_{50} (Fig. 3.2b) and v_{50} (Fig. 3.2c). Characteristic but not exclusive of HEWL aggregation, the low reproducibility of the results further prevents definitive conclusions to be taken without testing wider ranges of protein concentrations and numerous other replicates[142,145,149,150]. The previously reported formation of intermediate and off-pathway species during amyloid fibrillization of lysozyme should however produce atypical kinetic signatures and explain in part the poor reproducibility indexes. Parallel phase transition processes, such us the formation of insoluble oligomers and protein precipitates, acts as a sink of the soluble amyloid pool thus affecting the rate at which amyloid fibrils are formed and their final amount. Since the non-amyloidogenic pathways

also involve stochastic nucleation steps, their presence is expected to increase the overall variability of the results.

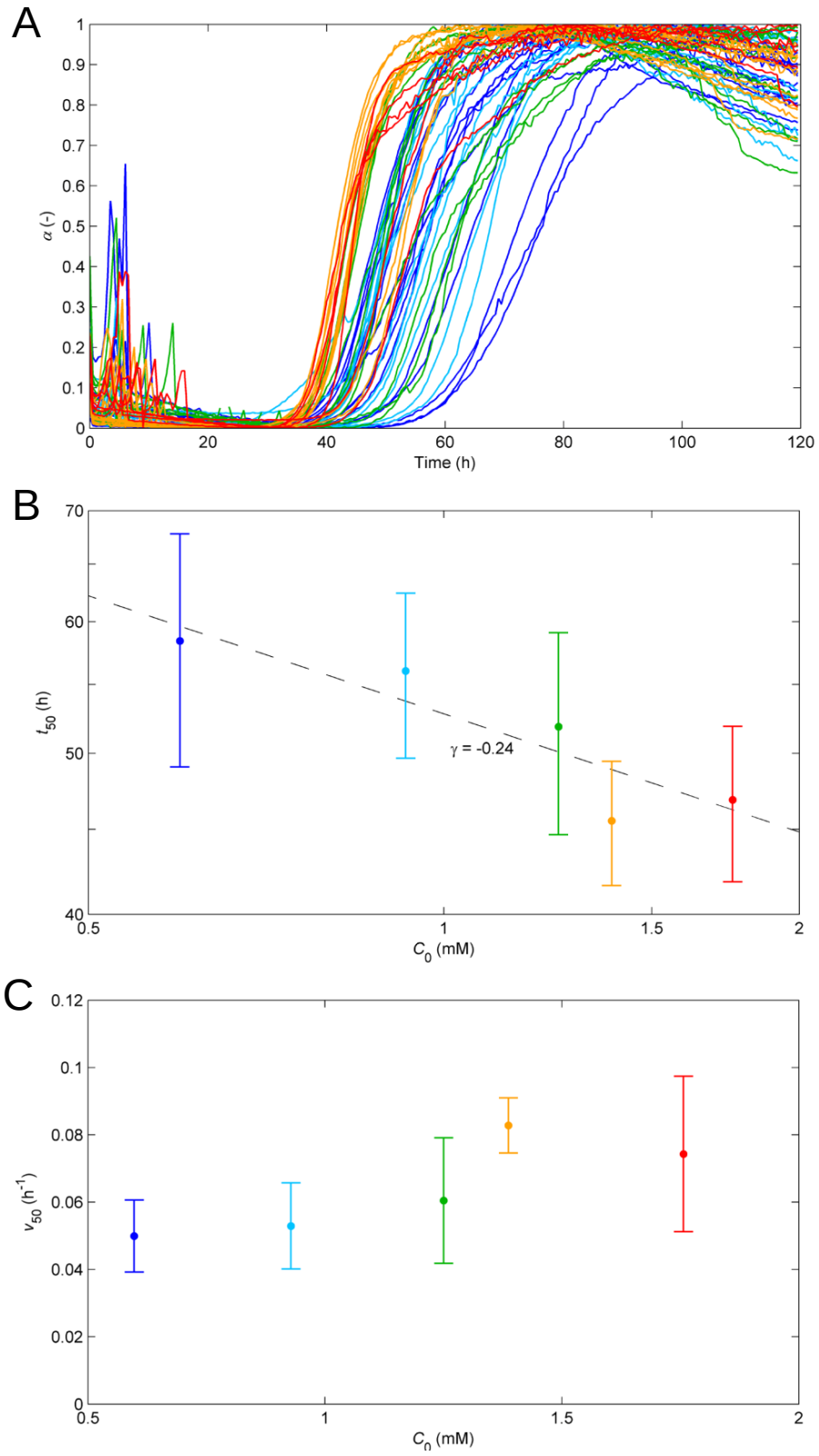


Figure 3.2. Influence of the initial protein concentration on the kinetics of amyloid fibril formation of HEWL at pH 1.6 and 60 ºC. (A) ThT fluorescence

increase during the aggregation of 0.60 mM (blue), 0.93 mM (light blue), 1.25 mM (green), 1.39 mM (orange) and 1.76 mM (red) HEWL represented in normalized units as a function of the incubation time. (B,C) Influence of the initial HEWL concentration (B) on the time required to reach 50 % completion (t_{50}) and (C) on the aggregation rate at the same instant (v_{50}). Log-log and linear-linear scales are adopted in (B) and (C), respectively; symbols and error bars represent mean values and standard deviations. (B) The linear fit (solid line) indicates an exponential scaling factor γ of −0.24.

The impact of an additional source of monomer depletion besides primary nucleation, secondary nucleation and fibril elongation was investigated within the CLM framework by introducing an off-pathway nucleation step, which, similar to amyloid nucleation, is considered to be second-order in relation to supersaturation. From this theoretical exercise, which is described in detail in Experimental Procedures, we found practical ways to identify supplementary kinetic steps from a single aggregation curve (Figs. 3.3 and 3.4); as shown in Figure 3.3A, the conventional hyperbolic/sigmoidal curves previously represented in Figure 3.1A are expected to show a noticeable linear phase when amyloid conversion is expressed as $\alpha/(1-\alpha)$ and represented as a function of time in a log-linear scale. Note that $\alpha/(1-\alpha)$ corresponds to the mass of fibrilsalready produced divided by the mass of fibrils still to be formed. The linear phase starts after elapsed a period of time of $\simeq k_a^{-1}$ and lasts until the end of the reaction. In the case of sigmoidal curves (low k_b values in Fig. 3.3A), this interval should include the whole fast growth period. As a first fingerprint of the presence of off-pathway species, a disturbance to the linear profile is shown in Figure 3.3B with the initial concave phase being prolonged until later stages of amyloid fibril formation. As the amyloid reaction approaches completion, the concave phase is immediately succeeded by a convex phase, whose presence might be difficult to identify in practice due to the increased signal noise of $\alpha/(1-\alpha)$ data. As Figures 3.4E and 3.4F also show, the deviations from linearity result from the existence of a parallel nucleation step characterized by the rate constant k_{off}. Representing the measured HEWL aggregation curves in the modified coordinates consistently show nonlinear concave trends during the fast growth stages - Figure 3.3C. Owing to fluorescence noise amplification, the initial and final reaction stages are omitted in Figure 3.3C. Log-linear representations of the measured $\alpha/(1-\alpha)$ with the incubation time are a new probe for the presence of off-pathway aggregates with facile implementation during high-throughput inhibitor screenings. However, and as discussed next, the information provided by

representations such as Figure 3.3C is essentially qualitative and requires additional scaling studies in order to be consolidated.

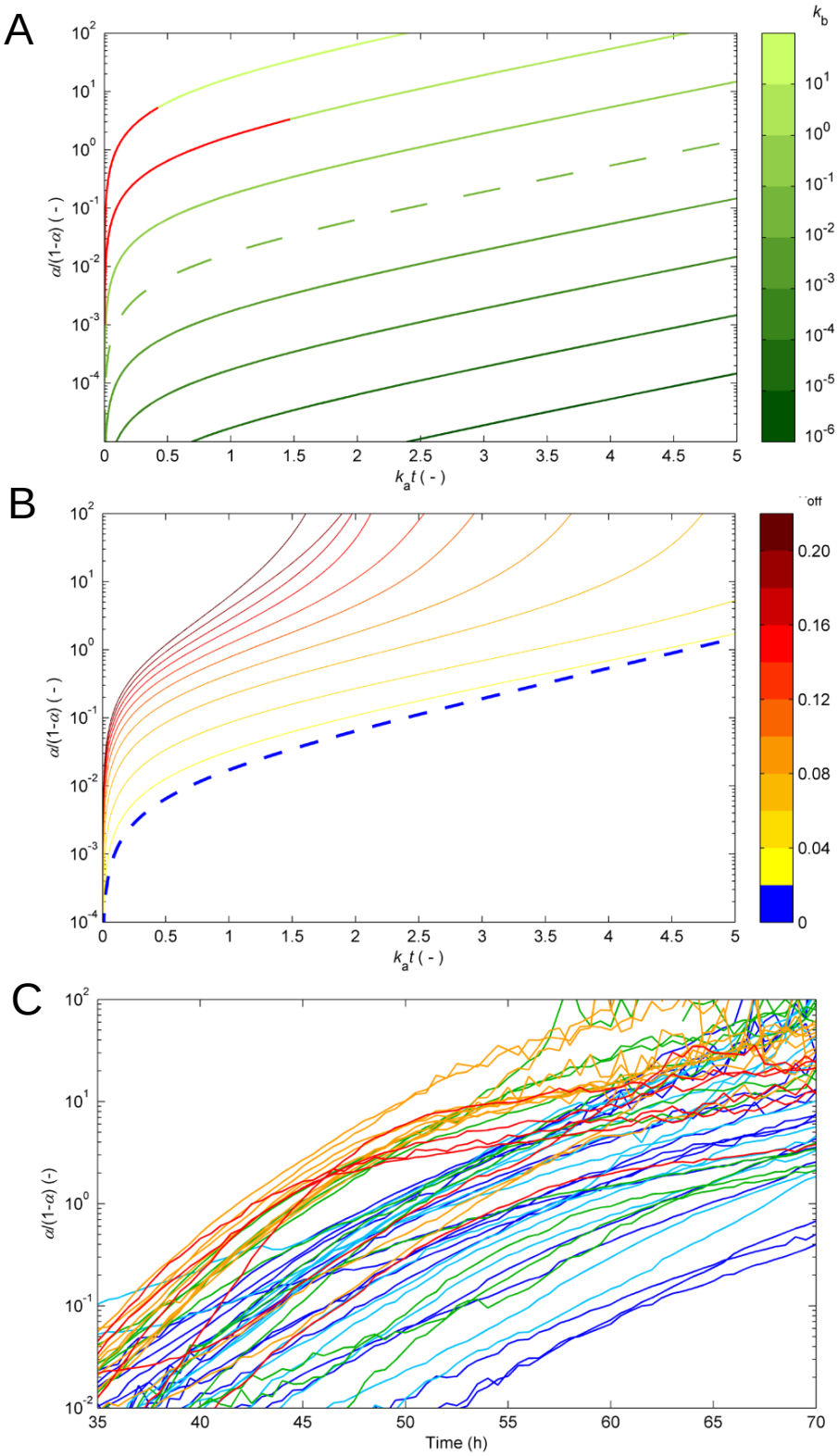


Figure 3.3. Kinetic signatures suggesting the presence of off-pathway species. (A) The typical CLM aggregation curves in Figure 3.1 (without OPA) are represented in modified log-linear coordinates relating $\alpha/(1-\alpha)$ vs. dimensionless

time (same color code as in Fig. 3.1A). Linear relationships are observed for $k_a t>1$ (after the lag phases in Fig. 1a are surpassed). (B) Effects provoked by other phase transition processes besides the nucleation and growth of amyloid fibrils. The linear phase rapidly vanishes in the presence of parallel nucleation events characterized by the rate constant k_{off}. Numerical results obtained from Simulation 1 (Experimental Procedures), using $k_b=10^{-2}$ and the values of k_{off} indicated in the color bar. Dashed lines in (A) and (B) correspond to the same result. (C) The measured aggregation curves of HEWL in Figure 3.2A are represented in the modified coordinates (same color code as in Fig. 3.2A). The concave phase prolonged until amyloid conversions close to 1 suggest the existence of OPA.

Ideal curves with marked linear phases as in Figure 3.3A do not necessarily mean the absence of supplementary pathways, which might take place at much slower rates than amyloid fibrillization. On the other hand, prolonged concave phases eventually followed by a convex phase (Figs. 3.3B, 3.3C, and Figs. 3.4E and 3.4F) are a necessary but not sufficient condition for the existence of parallel phase transition processes, seeing that similar outcomes might be produced by other phenomena, e.g., fluorescence quenching. The results obtained from Simulation 1 (Experimental Procedures) and represented in Figures 3.4A and 3.4B suggest additional evidences based on the variability of the endpoint signals F_∞ . In principle, these thermodynamically-determined measurements should be highly reproducible since they are not subject to stochastic contingencies as, for example, the nucleation steps.Figures 3.4A and 3.4B show, however, that the endpoint signals are expected to change as different values of k_b and k_{off} are considered. This is understandable in view of the existing competition between amyloid and off-pathway nucleation: the faster is one process relatively to the other, the larger is its share of the total soluble protein. In turn, if no OPA takes place, it does not matter thermodynamically whether amyloid formation is fast or slow, reproducible or changeable, because the value of F_∞ is mainly determined by the $\left(C_0 - C_a^*\right)$ difference. The experimental results previously shown in Figure 3.2A are associated to highly variable end-point fluorescence signals (data not shown). Such low reproducibility is normally associated to the nucleation rate constant k_b due to the exponentiation of the lag-time variability[6]. The propagation of the kinetic uncertainty to the F_∞ values can be justified by the occurrence of parallel nucleation processes, both characterized by fluctuating rate constants (k_b and k_{off}).

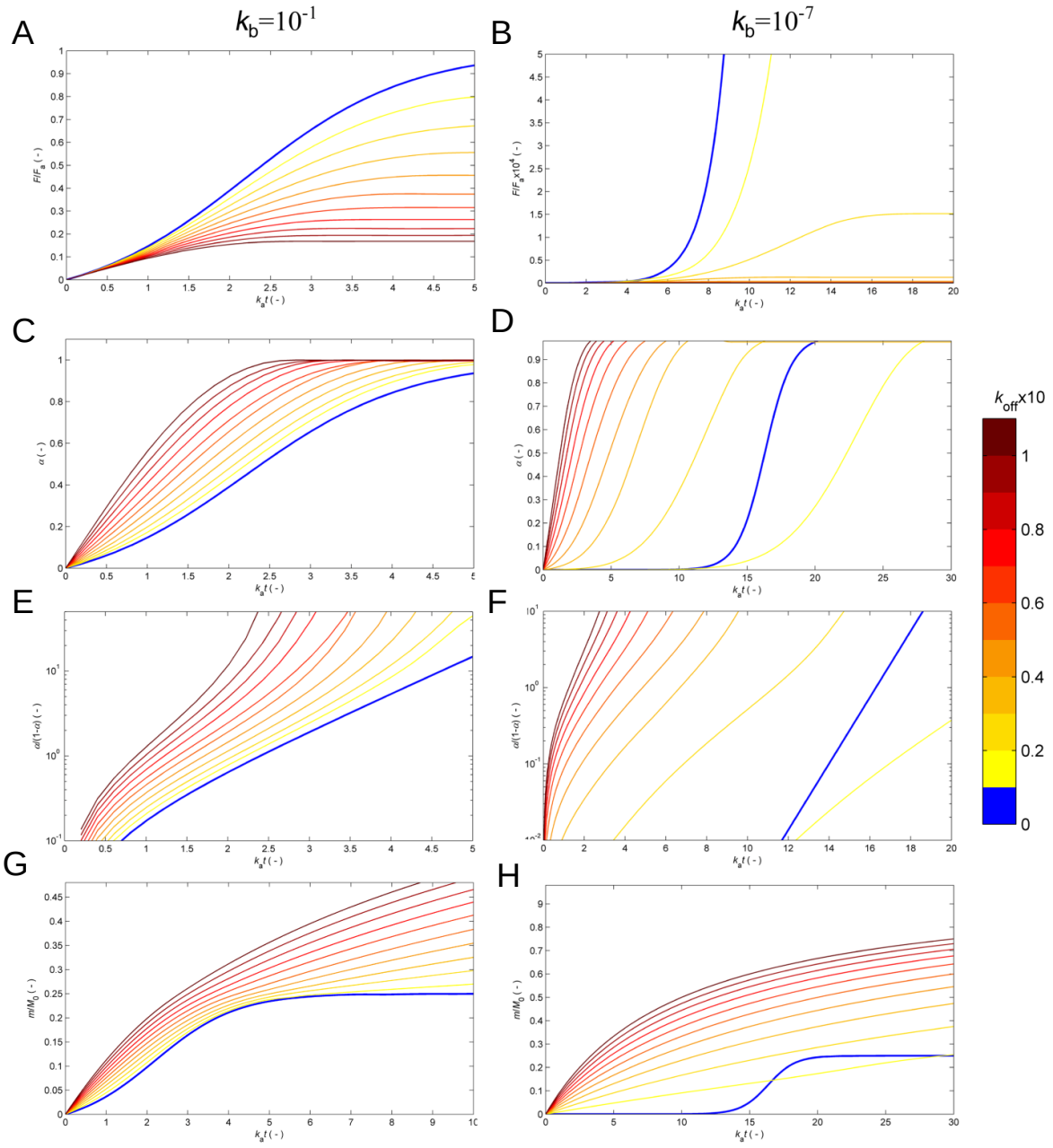


Figure 3.4. Numerical solutions of the extended CLM accounting off-pathway aggregation. Equation 3.10 was solved as described in Simulation 1 using k_b=10^{-1} (left panel) and k_b=10^{-7} (right panel) and the values of k_{off} given in the right side color bar. Results are expressed as the variation with the normalized time $k_a t$ of (A, B) the predicted amyloid fluorescence signal F/F_a (equivalent to α_a), (C, D) the amyloid conversion α, (E, F) the modified amyloid conversion $\alpha/(1-\alpha)$ in log-linear scale and (G, H) the dimensionless amount of total aggregates m/M_0.

OPA Scaling Laws:

After having scrutinized the progress curves of HEWL aggregation subject to different types of normalization and having analyzed the variability of the endpoint fluorescence signal, it remains to be discussed the peculiar scaling laws of t_{50} and v_{50} with

protein concentration. In fact, the exponent γ determined from the results in Figure 3.2B, and the aggregation rate data represented in Figure 2C are indicative of abnormally weak influence of C_0 on the two half-life coordinates. When only amyloid fibrils are formed, minimum values of $|\gamma| \simeq 1$ are predicted by the CLM under conditions of low nucleation rates (Fig. 3.1b), for which v_{50} also reduces to $\simeq 4k_a$, a linear function of C_0 [6]. The measured relationships shown in Figures 3.2B and 3.2C are unexpected not only by the 2-parameter CLM as by any other current model (see discussion of Fig. 3.1). Once again a solution to this problem seems possible by extending CLM to account the formation of non-amyloidogenic precipitates. The altered role of the initial protein concentration is studied in Simulation 2 (Experimental Procedures), with Figure 3.5 showing the numerical solutions assuming predominant OPA ($k_b/k_{off} \ll 1$). Figure 3.5A shows that the duration of the lag phase hardly changes with the protein concentration($\gamma \simeq 0$) and that t_{50} may even increase with C_0 ($\gamma > 0$). This apparently contradictory result is explained by the presence of a competitive off-pathway step that is comparatively more favored by higher C_0 values than the amyloid aggregation step. Equally, Figure 3.5B shows that the amyloid aggregation rate v_{50} can be weakly influenced by C_0 or even decrease as C_0 increases. To be observed, these paradoxical kinetic results require high values of k_{off} in order that the considered range of protein concentration is above the critical limit for amyloid formation. Figure 3.5 also shows sudden variations of the $t_{50}(C_0)$ and $v_{50}(C_0)$ scaling factors taking place along narrow ranges of C_0 values. In agreement with estimations of γ taken from literature[142], the broken curves are not confined to protein concentrations close to the amylois solubility (as in Fig. 3.1b) but can be observed for C_0/C_a^* values $\gg 1$. Our selection of k_b and k_{off} values in Figure 3.5 places the region of weak C_0 dependencies in the same C_0/C_a^* range as that used during the HEWL aggregation experiments. While reconciling the results of Figures 3.2B and 3.2C with the theory of protein aggregation, this agreement is the first step towards a univocal, all-inclusive validation of the model of which the following section presents a numerical attempt.

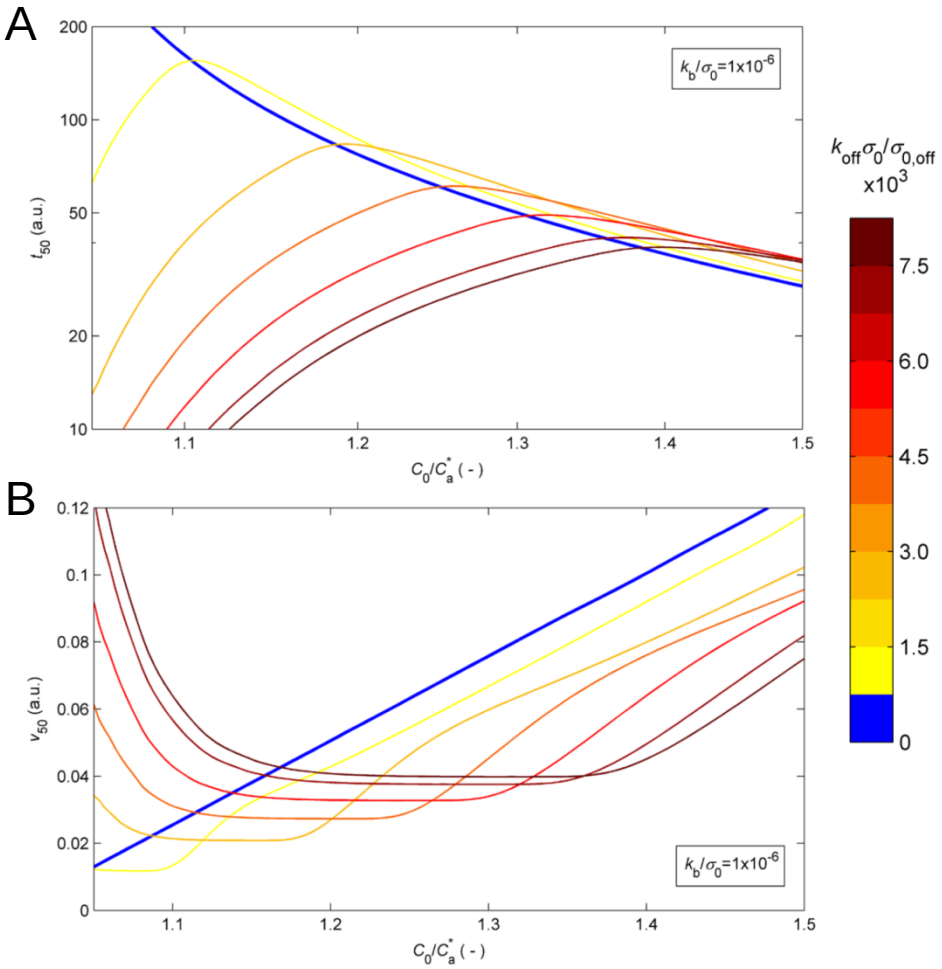


Figure 3.5. Influence of the protein concentration on the half-life coordinates when OPA is predominant. Solutions of the extended CLM calculated for $k_b/\sigma_0 = 10^{-6}$ given as (A) the log-log representation of t_{50} as a function of the C_0/C_a^* ratio and (*B*) the variation of v_{50} with C_0/C_a^*. (A,B) Different colors mean different C_0-independent k_{off} values as indicated by the color bar in the right side. Numerical details given in Simulation 2 (Experimental Procedures).

Protein depletion confirms OPA prevalence:

At this point, we have accumulated a set of evidences suggesting that amyloid fibrils are only the tip of an iceberg hiding a crowd of ThT-invisible aggregates. Although independent from each other, these evidences stem, in all cases, from conventional ThT fluorescence measurements. The first estimations of k_b values ~10^{-7}, indicate that the amyloid nucleation is many order of magnitudes slower than the autocatalytic steps. On the other hand, k_{off} values comprised between ~10^{-2} and ~10^{-1} configure a case where off-pathway species are produced at much higher ratesthan amyloid fibrils. These conclusions were further tested by complementary techniques, namely, by checking whether the

depletion of protein from solution matches the observed kinetics of amyloid fibril formation (Figs. 3.6a and 3.6b). By changing the focus from the amyloid fibrils to the dissolved protein we also wanted to measure the equilibrium concentration for long reaction times as an estimation of the thermodynamic solubility. Independent HEWL samples incubated at the same conditions of pH and temperature as during the fluorimetric assays were periodically filtered and analyzed spectrophotometrically at 280 nm (see Materials and Methods). The results in Figure 3.6A show that solute depletion starts immediately after incubation, i.e., before the formation of any detectable amount of amyloid fibrils (Fig. 3.6B); moreover, phase transition processes continue to occur many days after the plateau in fluorescence emission is reached. Besides confirming the predominance of OPA, these results also indicate that HEWL amyloid fibrils equilibrate with the solution earlier, and at higher protein concentration, than the other aggregates. The simulated curves in Figure 3.6A were computed using the estimates of k_b and k_{off} that followed from ThT fluorescence kinetic analysis and using the values of amyloid solubility (C_a^*) and off-pathway solubility (C_{off}^*) of 1.50 mM and 0.75 mM corresponding to estimations of HEWL concentration after ~90 h and ≫1000 h incubation, respectively. The calculated dimensionless time needed to obtain the same total conversion as after 700 h incubation was $k_a t = 100$, from which the value of $k_a = 1/7$ h^{-1} was determined. The good agreement between simulated and measured results reported in Figure 3.6A was achieved using a relatively narrow range of k_{off} values (color bar), meaning that the expectable variability of this parameter is able to explain the observed scattering of the measured profile. Despite all the converging evidences, the set of values chosen for k_a, k_b and k_{off} should be taken as approximate guesses rather than as definitive results. This is because the adopted equilibrium concentrations for amyloid (C_a^*) and off-pathway (C_{off}^*)aggregates are expected to differ substantially from the real thermodynamic solubilities. Not only the solutions are highly concentrated and nonideal, but also the chemical potential of the solute seems to be drastically influenced by the presence of aggregates in a process akin to volume exclusion effects[151]. The asymptotic HEWL concentration of 0.75 mM estimated for long reaction times implies that aggregation assays conducted at concentrations below this limit would not produce off-pathway precipitates, let alone amyloid fibrils. Yet, as the results in Figure 3.2 demonstrate, HEWL concentrations as low as 0.60 mM continue to produce aggregates that stain positive for ThT. The thermodynamic concentration of

dissolved protein seems to be decreased by the presence of aggregates, which, combined with changes in solution viscosity may explain why the phase transition processes cease at high HEWL concentrations and high content of total aggregates. These effects, to be discussed in detail elsewhere, imply a faster supersaturation decrease than that considered in our model. Therefore, they should also account for the differences observed in Figure 3.6B where the measured aggregation curves move faster towards equilibrium than the expected by the extended CLM for the selected set of parameters. Using real protein concentrations in our theoretical simulations would require time-evolving activity coefficients that are not available right now.

Complementary structural, morphological and size-distribution data:

The changes in the secondary structure of HEWL were followed by recording the far-UV spectra along the different phases of the aggregation process (Fig. 3.6c). The far-UV spectrum of non-incubated HEWL shows the 208 nm and 222 nm bands characteristic of the α-helical structure of the native protein[152]. Upon the first day of incubation, the ellipticity strongly decreases and the minimum CD intensity takes place at slightly lower wavelengths. The same tendency is observed until day 4, yet with smaller variations of CD. These results indicate the formation of pre-amyloid structures, presumably amorphous aggregates, which are less helical than the initial conformation. The formation of ThT-positive aggregates after the first day of incubation (Fig. 3.6b) did not provoke the position of minimum CD to clearly deviate from ~208 nm to ~215–220 nm as expected for β-sheet structures. Therefore, and as suggested by the protein depletion results, amorphous aggregation seems to prevail over amyloid fibrillization since the beginning of incubation. The CD intensity continues to decrease from day 4 to 7 as a consequence of the formation of off-pathway aggregates and the increased number and size of scattering objects in solution[131]. No amyloid fibrils are formed during this period as indicated by the ThT fluorescence plateau in Figure 3.6B.

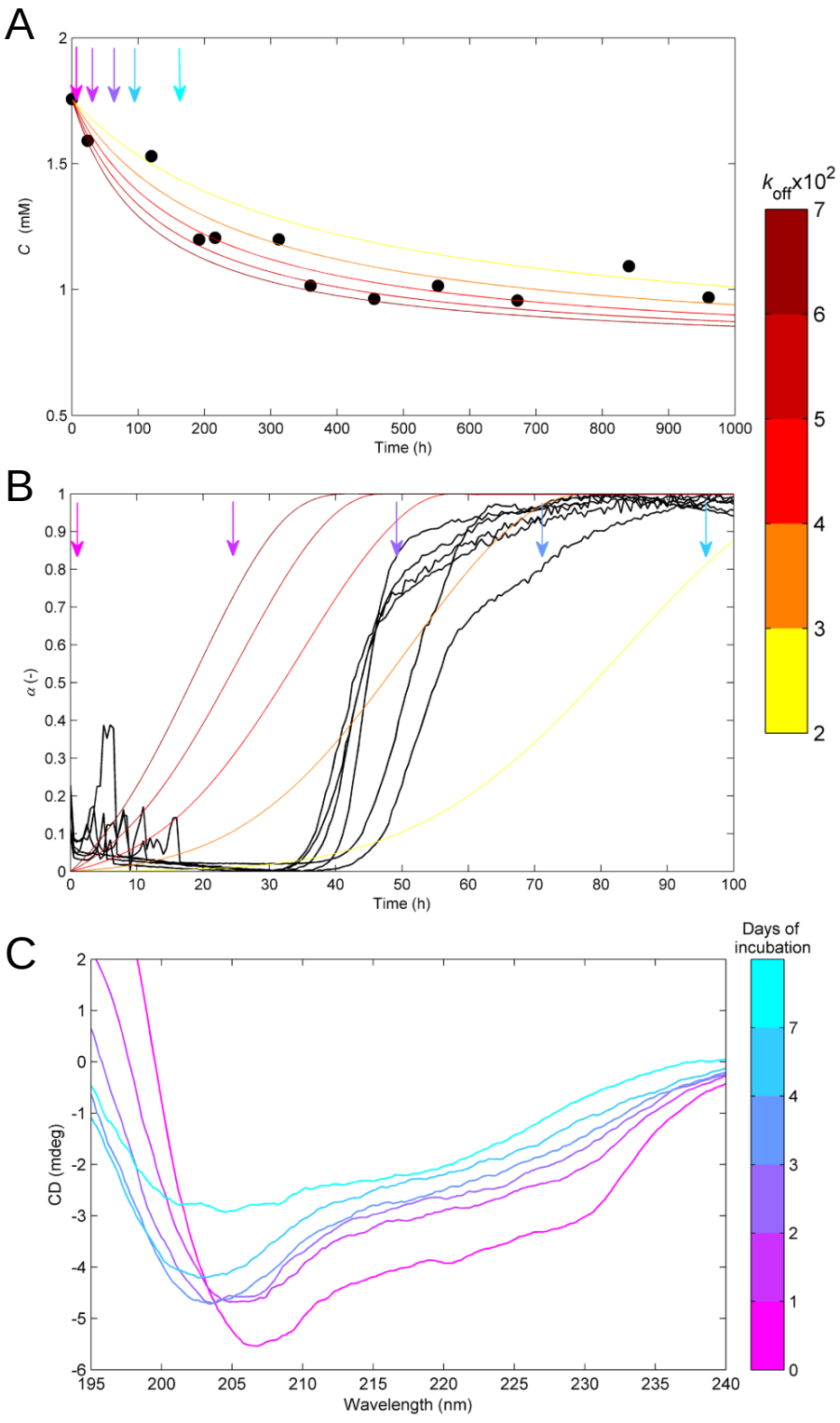


Figure 3.6. Complementary data suggesting predominant off-pathway aggregates over amyloid fibrils. (A) Symbols: depletion of soluble HEWL at pH 1.6 and 60 ºC with time measured by UV absorption at 280 nm after filtration through a sterile 0.22 µm filter. Lines: protein concentration decrease predicted by the extended CLM. (B) Amyloid fibrillization followed by ThT fluorescence (black lines) under the same conditions of pH, temperature and initial HEWL concentration (1.76 mM) as in (A). Different colored lines: same numerical solution as in (a) expressed as the amyloid content increase with time. HEWL depletion in (A) starts

before, and continues after the formation of amyloid fibrils in (B). (A,B) Numerical results obtained from Simulation 2 (Experiential Procedures), using k_a=1/7 $h^{-1,}$ k_b=10^{-7} and the range of k_{off} values indicated by the color bar in the right side; arrows: different instants of time where samples were analyzed using CD spectroscopy and DLS. (C) Far-UV CD spectra of HEWL samples incubated for 0, 1, 2, 3, 4 and 7 days measured at 20 ºC and at a 1:400 dilution. Major changes in the CD spectrum occur during the first day of incubation and from day 4 to 7, when no measurable amount of new amyloid fibrils is formed.

To gain insight into the size and morphology of the aggregates present in solution, samples incubated for different time periods were further analyzed using DLS and Atomic Force Microscopy (AFM) – Figure 3.7. The size distributions obtained from scattered light intensity measurements reveal how the populations of soluble protein and insoluble aggregates evolve with time (Fig. 3.7A). The relative weight of the peak centered near a hydrodynamic radius R_h of 2 nm is an indirect estimate of the amount of soluble HEWL. As shown in the inset of Figure 3.7A, the variation of this fraction during the incubation time resembles that of protein concentration measured by UV absorption in Figure 3.6A. This is another independent verification that phase transition processes occur before and after amyloid fibril formation. Even though particles with R_h greater than 50 nm are identified before incubation at 60 °C, the first DLS dataset is not sufficiently resolved to distinguish between the different pre-aggregation species. As the incubation starts, the distribution of particles centered near 70 nm becomes better defined simultaneously with the formation of the amorphous aggregates. Coinciding with the period of fast ThT fluorescence increase (Fig. 3.6b), a third peak centered near ~500 nm emerges at the end of day 2 as the likely result of the formation of mature amyloid fibrils. After that, the two aggregate peaks become blended in a single distribution centered in a hydrodynamic radius still close to 100 nm but with the right tail exceeding the submicron range. Since the scattering intensity of a particle is proportional to the sixth power of its diameter, the results in Figure 3.7A suggest the existence of a population of amorphous aggregates with mean R_h of 70 nm dominating over a second population of greater R_h, presumably mature amyloid fibrils, having a widely dispersed size distribution owing in part to fibril breakage. Giventhe limitations of DLS to morphologically describe less prevalent, non-spherical particles[153], the amyloidogenic samples were also characterized using AFM and Transmission Electron Microscopy (TEM). The selected AFM image in Figure 3.7B confirms that 1.76 mM HEWL solutions at pH 1.6 already contain amorphous aggregates before incubation at 60 ºC. Present in all analyzed samples, these precipitates generally

have the shape of a disk with height ~5 nm and diameter within the range 20-150 nm also obtained for R_h in Figure 3.7A. Figure 3.8 shows the results of further examinations using TEM, with small worm-like aggregates being identifiedin Figures3.8B and 3.8E. It remains unclear whether these protofibrils twist over themselves to form disk-shaped aggregates or continue to develop to form amyloid fibrils. In Figure 3.7C we surprisingly identified the presence of mature amyloid fibrils as long as 2 μm and ~10 nm diameter at the end of the first day of incubation. Despite their size, these fibrils are scarce enough to remain undetected during ThT fluorescence and DLS measurements. Therefore, amyloid fibrils seem to grow much faster than they are nucleated in agreement with the low value estimated for parameter k_b (~10^{-7}). The AFM images in Figures 3.7D and 3.7E show well differentiated curvilinear fibrils, whereas the TEM images additionally report straighter fibrils generally surrounded by amorphous aggregates – see Figures 3.8E-F. These variations are the likely result of different sample handling, which in the case of AFM involved 1:400 dilution and spin-coated deposition onto silicon wafers.

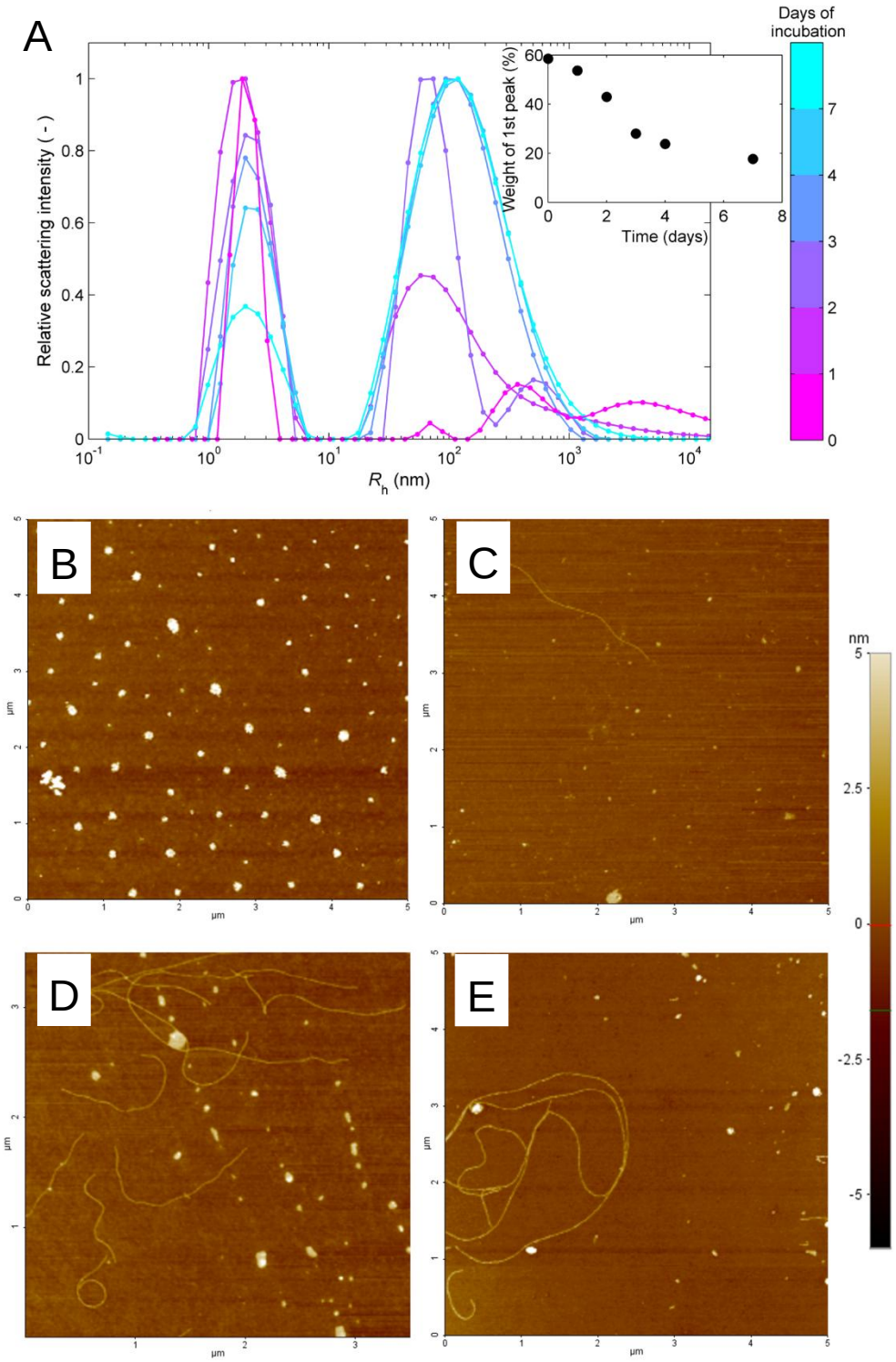


Figure 3.7. Aggregation of 1.76 mM HEWL at pH 1.6 and 60 ºC followed by DLS and AFM. (A) Distribution of hydrodynamic radio R_h obtained from DLS measurements at the different instants of time marked with an arrow in Figures 6A and 6B. Inset: the relative weight of the soluble HEWL peak centered near 2 nm decreases with time. (B-E) Morphology of HEWL aggregates observed with AFM after (B) 0, (C) 1, (D) 3 and (e) 4 days incubation; color bar in the right side: height scale common to all AFM images.

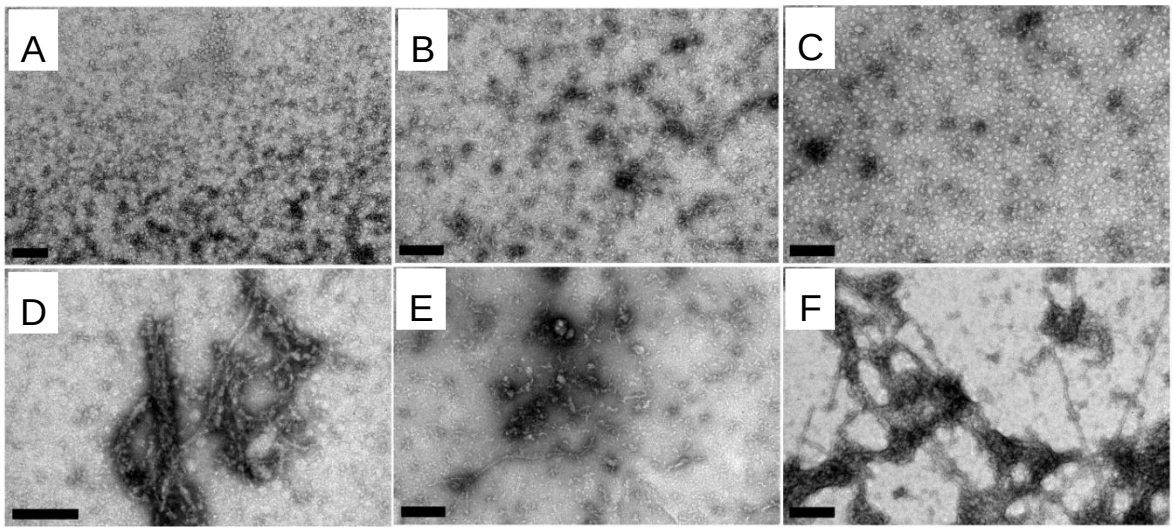


Figure 3.8. Aggregation of 1.76 mM HEWL at pH 1.6 and 60 ºC followed by Transmission Electron Microscopy. (A-C) Amorphous aggregates identified at the end of (A) 1, (B) 2 and (C) 3 days incubation; (D-F) amorphous aggregates and amyloid fibrils observed at the end of (D) 2, (E) 3 and (F) 4 days incubation. Scale bars represent 200 nm.

3.4 Discussion

Similarly to other phase transition phenomena, amyloid fibril formation takes place via a nucleation and growth mechanism until thermodynamic equilibrium is reached. The amyloid pathway may however comprise parallel steps and intermediate species that are no less relevant for the development of amyloidosis and neurodegenerative diseases than the deposition of fibrils itself. Protein aggregation curves measured *in vitro* using amyloid-specific dyes are shown to provide much more information about OPA than what is conventionally extracted. The set of canonical behaviors summarized in Figure 3.1 can be expected when off-pathway species are eitherabsent or present in minor amounts. Even if some of those kinetic results are not predicted by established theories, they are in conformity with the classical mechanism involving primary nucleation and auto-catalytic steps. The CLM was extended to include a third rate constant (k_{off}) characterizing the parallel nucleation step. The numerical solutions of the three-parameter CLM unveil the kinetic signatures characteristic of OPA; these include non-linear variation of $\alpha/(1-\alpha)$ with time when plotted in a log-linear scale, highly variable endpoint signals, values of the t_{50} absolute scaling factor $|\gamma|$ below 1 and sublinear increase of v_{50} with the protein

concentration. In cases of extensive OPA, increasing the protein concentration may even prolong the duration of the lag phase and decrease the amyloid aggregation rate. This angle of approach is, to our knowledge, totally original as the main focus has been on the detection and morphological/ toxicological characterization of the amorphous precipitates[121,123–125,129]. The new possibility of estimating the relative amounts of off-pathway and amyloid aggregates accentuates the need for cataloguing the deleterious species in each disease. After deciding what aspect of protein aggregation one wants to target, extensive screenings of off-pathway modulators can from now on be routinely implemented.

As a proof-of-concept the modified CLM was tested against ThT aggregation data of HEWL measured under conditions of low pH (1.6) and high temperature (60 ºC) known to produce off-pathway aggregates. The estimated values of k_a , k_b and k_{off} suggest that off-pathway species are produced much earlier and at higher rates than amyloid fibrils, which, nevertheless, rapidly reach their mature size once the stable nucleus is formed. This was supported by complementary analysis ofsoluble protein depletion with time, far-UV CD spectra, particle size distributions measured using DLS and aggregates morphology observed using AFM and TEM. We further concluded that amorphous aggregates generally have the shape of a disk with diameters comprised between 20 and 150 nm, even though small worm-like aggregates were also identified. Amyloid fibrils with 10 nm diameter are consistently longerthan 1 μm and show a broad size distribution as a likely consequence of fibril breakage.

While offering an explanation for puzzling kinetic behaviors, our study contributes to a better understanding of the molecular basis of amyloid diseases and is expected to find practical application in neurodegenerative drug research. By using the analytical probes for OPA here proposed, libraries of small molecule compounds can be screened targeting the formation of amorphous aggregates without any additionalmeans being required besides the current amyloid-specific markers and high-throughput methods. Whether the kinetic signatures are attenuated or intensified by the screened molecules provides a valuable indication of their potential as off-pathway modulators. This therapeutic strategy aims at inhibiting or promoting OPA according to the disruptive or stabilizing effect that non-fibrillar species may have on the regulatory mechanisms during disease.

Chapter 4

Macromolecular crowding effects and OPA

The availability of soluble protein is decreased in amyloidogenic media crowded with insoluble off-pathway aggregates

Rosa Crespo[1], Eva Villar-Alvarez[2], Pablo Taboada[2], Fernando A. Rocha[1], Ana M. Damas[3] and Pedro M. Martins[1,3]

[1]From LEPABE, Laboratório de Engenharia de Processos, Ambiente, Biotecnologia e Energia, Departamento de Engenharia Química, Faculdade de Engenharia da Universidade do Porto. Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

[2]Departamento de Física de la Materia Condensada, Facultad de Física, Universidad de Santiago de Compostela, Spain

[3]ICBAS – Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto, Portugal

To whom correspondence should be addressed: Pedro M. Martins, ICBAS – Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Rua de Jorge Viterbo Ferreira n.º 228, 4050-313 Porto, Portugal, Tel.: (+351) 220428137; E-Mail: pmmartins@icbas.up.pt

Summary

The screening of drugs candidates for the treatment of amyloidosis and neurodegenerative diseases frequently involves in vitro measurements of amyloid fibril formation. Macromolecular crowding and off-pathway aggregation (OPA) are, by different reasons, two important phenomena affecting the scalability of amyloid inhibitors and their successful application in vivo. On the one hand, the cellular milieu is crowded with macromolecules that drastically increase the effective (thermodynamic) concentration of the amyloidogenic protein. On the other, off-pathway aggregates, rather than amyloid fibrils, are increasingly appointed as the causative agents of toxicity. The present contribution reveals that insoluble off-pathway aggregates of hen egg-white lysozyme (HEWL) are a peculiar type of crowding agents that, unlike classical macromolecular crowders, decrease the thermodynamic concentration of protein. Illustrating this effect, OPA is shown to resume after lowering the fraction of insoluble aggregates at constant soluble HEWL concentration. Protein depletion and thioflavin-T fluorescence progress curves indicate that OPA rebirth is not accompanied by additional amyloid fibril formation. The crystallization-like model extended to account for OPA and time-dependent activity coefficients is able to fit multiple kinetic results using a single set of three parameters describing amyloid nucleation, autocatalytic growth and off-pathway nucleation. The list of fitted results notably includes the cases of aggregation rebirth and all types of progress curves measured for different HEWL concentrations. The quantitative challenges posed by macromolecular crowding and OPA find here a unified response with broader implications for the efficient design of on- and off-pathway inhibitors.

4.1 Introduction

Macromolecular crowding and off-pathway aggregation (OPA) are two subjects under intense, yet mostly unlinked, research attention. Both phenomena play an important role on the pathogenesis of amyloidosis and neurodegenerative diseases characterized by the accumulation of amyloid deposits in organs and tissues. Fluid media crowded with macromolecules promote the formation of amyloid fibrils owing to volume-excluding effects consisting of steric interactions that may affect the native protein conformation[154,155], increase the effective protein concentration[151,156] and interfere with the diffusive behavior of macromolecules[157]. In depth analysis of these effects involves adding high concentrations of inert macromolecules to in vitro systems in order to reproduce the crowded environment of most biological fluids, whether intracellular or extracellular[154,158]. The discovery in 2012 that the crowding agent Ficoll-70 and its monomer, sucrose had approximately the same protein stabilizing effect contributed to extend the range of excluded volume effects to include non-specific soft interactions[159,160]. Off-pathway aggregates are hereby investigated as crowding agents during hen egg-white lysozyme (HEWL) aggregation. This aspect of OPA remains, to our knowledge, totally unexplored as the main focus has been on the detection and morphological/toxicological characterization of the non-amyloidogenic species[121,129,161,162]. By 'off-pathway aggregates' we are considering insoluble, high-order oligomers and precipitates[163–165], and we are excluding soluble oligomers and insoluble protofibrils whose intermediate role in amyloid assembly is not totally defined[166–168].

The association between the two seemingly unconnected subjects stems from our effort to quantitatively characterize OPA from amyloid fibrillization kinetics[169]. We realized that off-pathway aggregates not only deplete the soluble amyloid pool via an additional kinetic step, as they decrease the effective protein concentration by means of thermodynamic effects that needed characterization. A similar objective has previously been faced while modeling macromolecular crowders as modifiers of the activity coefficient value γ relating the overall density of molecules (or apparent concentration) C with the thermodynamic activity or effective concentration a $(=\gamma C)$ [169,170]. Existing statistical mechanical theories use necessarily simple descriptions of crowding agents as hard-spheres or hard-rods in a practical alternative to computational simulation methods of molecular dynamics[171]. Experimental validation of these theories often relies on indirect

measurements of amyloid fibrillization kinetics in the presence of model crowders, such as Dextran or Ficoll[170,171]. In dealing with activity coefficients, we suggested that protein concentrations measured after long incubation times (C_∞) are a closer indicator of the thermodynamic equilibrium that takes place when the chemical potential of the solution equals that of the insoluble fraction[13]. In absence of OPA, C_∞ is equivalent to the amyloid apparent solubility C_a^*, which in turn, should correspond to the amyloid activity a^* under very dilute, ideal solutions. Enthalpic disturbances involving, e.g., macromolecular crowders, off-pathway aggregates or hydrophobic interfaces do not alter the intrinsic value of a^* but are likely to change the apparent solubility C_a^* and, consequently, the value of the activity coefficient under saturated conditions γ_a^*. In the presence of OPA, the steady-state concentration C_∞ may as well differ from C_a^* in the cases where off-pathway aggregates are less soluble than amyloid fibrils, for which C_∞ corresponds to the off-pathway solubility C_{off}^* [126]. Therefore, besides competing kinetically with amyloid fibrils for soluble protein[13], off-pathway aggregates may have an additional thermodynamic role either disrupting or stabilizing the regulatory mechanisms during disease[126].

Different solubility values reflect the structural stability of the different aggregation states and determine, in part, their relative abundance in the cell[172]. Another important part is determined by kinetic barriers for nucleation and growth[1], whose relative magnitudes were recently determined for HEWL aggregation with identification of the main signatures left by the presence of OPA; namely, amyloid conversion α expressed as $\alpha/(1-\alpha)$ and represented as a function of time in semi-log plots becomes highly non-linear, whereas the half-life coordinates of time and aggregation rate become weakly dependent on the initial protein concentration C_0 [6]. While important for the development of a new class of off-pathway regulators, this analysis needed further thermodynamic insights in view of the holistic quantitative representation of amyloid aggregation. The present contribution provides an important piece of this puzzle thus helping to close the gap between in vitro results and their physiological counterparts during drug discovery.

4.2 Experimental Procedures

Protein Preparation:

HEWL powder obtained from Merk KGaA (Darmstadt, Germany) was dissolved in 25 mM HCl pH 1.6 and dialyzed against 25 mM HCl, pH 1.6, using a cut-off 3500 Da membrane (Spectrum, Fisher Scientific). The concentration of the dialyzed protein solution was determined by absorbance measurements at 280 nm using an extinction coefficient of 37752 $M^{-1}cm^{-1}$. Protein stocks were stored at 4 ºC for no longer than one week.

Protein Depletion Experiments:

The concentration of soluble HEWL was followed spectrophotometrically over time under amyloidogenic conditions reproducing the Thioflavin-T (ThT) fluorescence experiments previously carried out by us[167]. Five series of independent 1.0ml samples of HEWL with final concentrations of 0.60, 0.93, 1.25, 1.39 and 1.76 mM were incubated at 60ºC. For each concentration series, one sample at a time was periodically taken from the incubator and the suspension was filtered through a sterile 0.22 μm pore size PES membrane filter (Jet Biofil). After 1 hour at room temperature, the filtrate was diluted 1:50 in 25 mM HCl pH 1.6 and analyzed at 280 nm using 1 cm pathlength quartz cuvettes (Hellma GmbH & Co. KG, Müllheim, Germany) and an extinction coefficient of 37752 $M^{-1}cm^{-1}$.

Protein Depletion with Expanded Mixture Volume:

The effect of the fraction of aggregates on the rate and extension of protein depletion was investigated by expanding the reaction volume at constant soluble HEWL concentration, temperature and pH. The mixture volume of some of the samples used in protein depletion experiments was expanded from 1.0 ml to 2.0 ml after elapsed a variable period time (of 240, 456, 456, 360 and 456 h for the 0.60, 0.93, 1.25, 1.39 and 1.76 mM samples, respectively). Thermostatized HEWL solutions freshly prepared to the same soluble protein concentration as that of the incubated samples (0.52, 0.67, 0.70, 0.91 and 0.96 mM for the 0.60, 0.93, 1.25, 1.39 and 1.76 mM samples, respectively) were used. Samples with and without mixture volume expansion continued incubation and were subject to periodic soluble HEWL concentration measurements as described for protein depletion experiments.

Dynamic Ligth Scattering (DLS):

DLS measurements were performed using an ALV/DLS/ SLS-5000F, SP-86 goniometer system (ALV-GmbH, Langen, Germany) as previously described[167]. The previous analysis of six independent 1.76 mM HEWL samples incubated for different periods of time at 60 ºC and pH 1.6[167] was complemented with new measurements using freshly prepared 3.57 mM HEWL samples in 0.2 M sodium acetate buffer, pH 4.7. Measurements were made at least three times at a scattering angle 90º to the incident beam for 5-10 min.

Activity coefficients:

The results obtained during protein depletion experiments with and without mixture volume expansion showed that the apparent HEWL concentration estimated for long incubation times C_∞ is a function of the fractional volume of aggregates, with lower values of C_∞ obtained for lower insoluble aggregates content. The different values of C_∞ should, however, correspond to the same thermodynamic activity a^* of the less soluble, off-pathway aggregates. This activity was approximated to the limit C_∞ value of 0.40 mM obtained for the less concentrated HEWL samples after the volume expansion step since these are the conditions originating the lowest fraction of insoluble aggregates in solution. The activity coefficients under saturated conditions γ_{off}^* were estimated dividing the fixed value of a^* by the limit C_∞ values obtained for each protein depletion experiment carried out without mixture volume expansion.

Mathematical Model Derivation:

In its simplest version, the crystallization-like model (CLM) has a closed-form solution describing how the molecular mechanisms of nucleation and growth affect the evolution of amyloid conversion (α) with time (t) during seeded or unseeded reactions[6,130]. The 2-parameter CLM equation includes the autocatalytic rate constant k_a accounting for the fibril elongation and secondary nucleation phenomena[6,167], and the rate constant k_b indicative of the relative magnitude of primary nucleation over autocatalytic steps. We have recently extended the CLM to include OPA by introducing a parallel nucleation step with normalized rate constant k_{off} [167]. Similarly to amyloid homogeneous nucleation, this step is considered to be second-order in relation to the thermodynamic

driving force for phase transition represented by supersaturation S. As time-dependent values of the activity coefficient γ are not easily obtainable, the primary definitions of S_a and S_{off}:

$$S_i = \frac{\gamma C - \gamma_i^* C_i^*}{\gamma_i^* C_i^*} \qquad (4.1)$$

have been approximated by the a relative concentration differences defined as

$$\sigma_i = \frac{C - C_i^*}{C_i^*} \qquad (4.2)$$

for either amyloid aggregation (subscript $i = a$) or OPA (subscript $i = off$). As a consequence of this approximation, rate constants fitted to kinetic data are only apparent estimators of the real event frequency of nucleation and growth. The differential form of the 3-parameter model equations reads:

$$\frac{\mathrm{d}\alpha_a}{\mathrm{d}t} = k_a \frac{\sigma_a}{\sigma_{a,0}} \left(k_b \frac{\sigma_a}{\sigma_{a,0}} + \alpha_a \right) \qquad (4.3a)$$

$$\frac{dm}{dt} = k_a \frac{\sigma_a}{\sigma_{a,0}} M_a \left(k_b \frac{\sigma_a}{\sigma_{a,0}} + \alpha_a \right) + k_{off}^* V \sigma_{off}^2 \qquad (4.3b)$$

where α_a corresponds to the instantaneous amount of amyloid fibrils (m_a) over the total amount expected in the absence of OPA (M_a); m is the instantaneous amount of total aggregates; k_{off}^* is the absolute rate constant for OPA; and V is the mixture volume. The definition of $\alpha_a = m_a / M_a$ implies that final value of 1 is only reached when no off-pathway aggregates are formed. In all other occasions, the value of α_a measured for long incubation times ($\alpha_{a,\infty}$) represents the amount of amyloid fibrils really produced over the maximum amount expected in the absence of OPA. Therefore, the conventional amyloid conversion α, computed, e.g., from the ThT fluorescence signal and comprised between 0 and 1, is equivalent to $\alpha = \alpha_a / \alpha_{a,\infty}$. If the instantaneous concentration of soluble protein C is known, the instantaneous conversion in *total* aggregates directly follows from mass balance to obtain:

$$\beta = \frac{m}{M_0} = \frac{C_0 - C}{C_0 - C_\infty} \qquad (4.4)$$

with M_0 being the final amount of total aggregates. The concentration of protein for long reaction times C_∞ corresponds to the apparent solubility of the less soluble species, which, in the case of HEWL, are the off-pathway aggregates (i.e., $C_\infty = C_{off}^*$) [167]. Equally, if the instant of amyloid saturation is known (using, e.g., amyloid-specific dyes), then the apparent amyloid solubility C_a^* corresponds to the value of C at that instant. The M_a/M_0 ratio is then computed as:

$$\frac{M_a}{M_0} = \frac{C_0 - C_a^*}{C_0 - C_\infty} \qquad (4.5)$$

In the present development of the CLM, activity-based supersaturation levels S are used in the model equations to replace the less realistic σ values. This involves using Eqs. 4.4 and 4.5 to express Eq. 4.1 as a function of the integration variable β for the cases of (i) amyloid aggregation,

$$\frac{S_a}{\sigma_{a,0}} = \frac{\gamma}{\gamma_a^*}\left(1 - \beta\frac{M_0}{M_a}\right) + \frac{1}{\sigma_{a,0}}\left(\frac{\gamma}{\gamma_a^*} - 1\right) \quad (4.6a)$$

and (ii) OPA,

$$\frac{S_{off}}{\sigma_{off,0}} = \frac{\gamma}{\gamma_{off}^*}(1-\beta) + \frac{1}{\sigma_{off,0}}\left(\frac{\gamma}{\gamma_{off}^*} - 1\right) \quad (4.6b)$$

Accordingly, Eqs. 4.3a and 4.3b can now be expressed in terms of activity-based supersaturation levels S, conversion β and dimensionless time $\theta = k_a t$,

$$\frac{\mathrm{d}\alpha_a}{\mathrm{d}\theta} = \frac{S_a}{\sigma_{a,0}}\left(k_b\frac{S_a}{\sigma_{a,0}} + \alpha_a\right) \qquad (4.7a)$$

$$\frac{d\beta}{d\theta} = \frac{S_a}{\sigma_{a,0}}\frac{M_a}{M_0}\left(k_b\frac{S_a}{\sigma_{a,0}} + \alpha_a\right) + k_{off}\left(\frac{S_{off}}{\sigma_{off,0}}\right)^2 \quad (4.7b)$$

using the $S_a/\sigma_{a,0}$ and $S_{off}/\sigma_{off,0}$ definitions given by Eqs.4.6a and 4.6b, respectively, and the definition of M_a/M_0 given by Eq. 4.5; the definition of k_{off} follows from the direct dependence of k_a on $\sigma_{a,0}$ ($k_a = \kappa\sigma_{a,0}$) [6] and from the alternative definition of the total amount of aggregates as $M_0 = (C_0 - C_\infty)M_w V$:

$$k_{off} = \frac{k_{off}^{*} V \sigma_{off,0}^{2}}{k_a M_0} = \frac{k_{off}^{*}}{M_w \kappa}\left(\frac{\sigma_{off,0}}{\sigma_{a,0}}\right) \qquad (4.8)$$

where M_w is the molecular weight of the amyloidogenic protein.

Numerical Methods:

Equations 4.7a and 4.7b were numerically solved using Matlab® subject to the initial conditions $\alpha_a(0) = 0$ and $\beta(0) = 0$ and testing the set of rate constants k_a, k_b, k_{off} that best described the $\alpha(t)$ and $\beta(t)$ progress curves measured during ThT fluorescence and protein depletion experiments, respectively. Other unknowns besides the fitting parameters include values of σ_0 and M_a/M_0 that can be computed from Eqs. 4.2 and 4.5 using the different initial HEWL concentrations C_0, and the values of C_a^* and C_∞ obtained from the protein depletion experiments. Amyloid fibrils are not considered to dissolve once the thermodynamic concentration of soluble protein decreases below amyloid solubility. After that limit, amyloid supersaturation S_a is considered 0. The time-dependent activity coefficients were estimated as described above using the linear relationship between γ and the concentration of insoluble aggregates (Table 4.1 and inset of Figure 4.1). Numerical fitting was initially carried out adopting k_a-normalized units of time (θ) thus reducing the fitted parameters to the values of k_b and k_{off}, which determine the amyloid and off-pathway timescales. After the relative timescales matched the observed ones, k_a was directly calculated as the θ/t ratio using the values of θ (simulated) and t (measured) that corresponded to the same total conversions β. The theoretical results were represented by the normalized amyloid conversion profiles $\alpha(t) = \alpha_a(t)/\alpha_{a,\infty}$, and by the concentration depletion profiles $C(t)$ computed from $\beta(t)$ using Eq.4.4. In a supplementary validation test, the fitted model was used to simulate the mixture volume expansion experiments carried out at constant HEWL concentration, temperature and pH; for that, Eqs.4.7a and 4.7b were solved subject to modified initial conditions corresponding to the amyloid conversion, total conversion and soluble HEWL concentration (α_a, β and C, respectively) predicted by the numerical simulations for the instants at which the mixture volumes were expanded. From that point on, the concentration of total aggregates used to calculate the activity coefficients were multiplied by a factor of $1/2$ (to account for the

dilution of the insoluble phase), and the new values of C_0 and C_∞ were used to compute M_a/M_0 and $\sigma_{off,0}$.

4.3 Results

'Solvation-in' effects:

Our practical approach to quantify OPA-induced thermodynamic effects is based on the concept of thermodynamic solubility as the activity of the solution phase equilibrating that of the solid phase[82]. If present during HEWL aggregation, enthalpic effects are expected to influence the equilibrium concentration of soluble protein estimated for long incubation times C_∞ but not the intrinsic thermodynamic activity a^* of the protein molecules at the aggregate-solution interface[6,167]. Table 4.1 shows that higher values of C_∞ are obtained as the initial HEWL concentration increases indicating that the final concentrations correspond to apparent solubility values. Consequently, the final amount of insoluble aggregates expressed as $(C_0 - C_\infty)$ did not increase with C_0 as much as it would be expected in the case of constant C_∞. As more insoluble aggregates are present, higher concentrations of HEWL molecules appear to be needed in order to achieve the same thermodynamic activity at the end of the assay. Therefore, and contrarily to the classical macromolecular crowders, HEWL aggregates decrease the effective protein concentration in solution. While promoting protein solubilization, this effect is associated to an increase in the solvent accessible surface of HEWL molecules[93]. The type of interactions here involved is certainly not limited to the van der Waals forces and repulsive electrical double-layer forces of the classical electrostatic theories, but should belong to the myriad of dispersion forces thought affect biological colloid systems[173]. This is especially so for the highly non-ideal amyloidogenic environment where hydration forces and quantum forces are interplaying in a ion-specific manner[173]. We borrow the concepts of salting-in and salting-out used to describe the action of salts[93] to generically name as 'solvation-in' and 'solvation-out' effects the increased/decreased protein solubility caused by the presence of protein aggregates and other colloids in solution.

Table 4.1. Protein depletion experiments carried out for different initial HEWL concentrations C_0 provide estimations of the saturation HEWL concentration C_∞, final amount of total aggregates $(C_0 - C_\infty)$ and apparent amyloid solubility C_a^*. The value of C_∞ is the apparent off-pathway solubility from which the activity coefficient γ_{off}^* is estimated in relation to a reference solubility value (0.40 mM) estimated for low aggregate content.

C_0 (mM)	C_∞ (mM)	$(C_0 - C_\infty)$ (mM)	C_a^* (mM)	γ_{off}^* (-)
0.60	0.45	0.15	0.50	0.89
0.93	0.45	0.48	0.71	0.89
1.25	0.50	0.75	0.92	0.80
1.39	0.55	0.79	1.10	0.67
1.76	0.70	1.06	1.30	0.57

Changing the focus from total aggregates to amyloid fibrils, we look now to the stages of aggregation coinciding with the final plateau of ThT fluorescence experiments[167]. The fibril-solution equilibrium required incubation periods no longer than 90 h, after which the apparent amyloid solubility values C_a^* could be estimated (Table 4.1). The existence of solvation-in effects is again suggested by increasing C_a^* values as the initial HEWL concentration increases. This variation partially explains the weak C_0 dependence of the end-point ThT fluorescence F_∞ described in Figure 4.1A: in principle, F_∞ should increase linearly with C_0 as the total amount of amyloid fibrils is determined by the $(C_0 - C_a^*)$ difference (red line). Since in practice C_a^* also increases with C_0, the expected end-point fluorescence signal becomes a weaker function of the initial HEWL concentration (blue line). Despite the improvement, the new dependency is not sufficient to reproduce the nearly invariant F_∞ values obtained for HEWL concentrations comprised between 0.60mM and 1.76 mM (symbols). As demonstrated next, the kinetic effect produced by OPA not only provides the other part of the explanation as it rationalizes the high variability associated to the end-point signal (error bars). This is to be illustrated within the

CLM framework taking into account self-crowding effects in the definition of the thermodynamic driving force for aggregation. As the phase transition evolves towards increasing amounts of total aggregates, activity coefficient values (γ) progressively smaller than 1 are utilized in Eq.4.1 in accordance with the linear relationship obtained under saturation conditions (Figure 4.1A, inset).

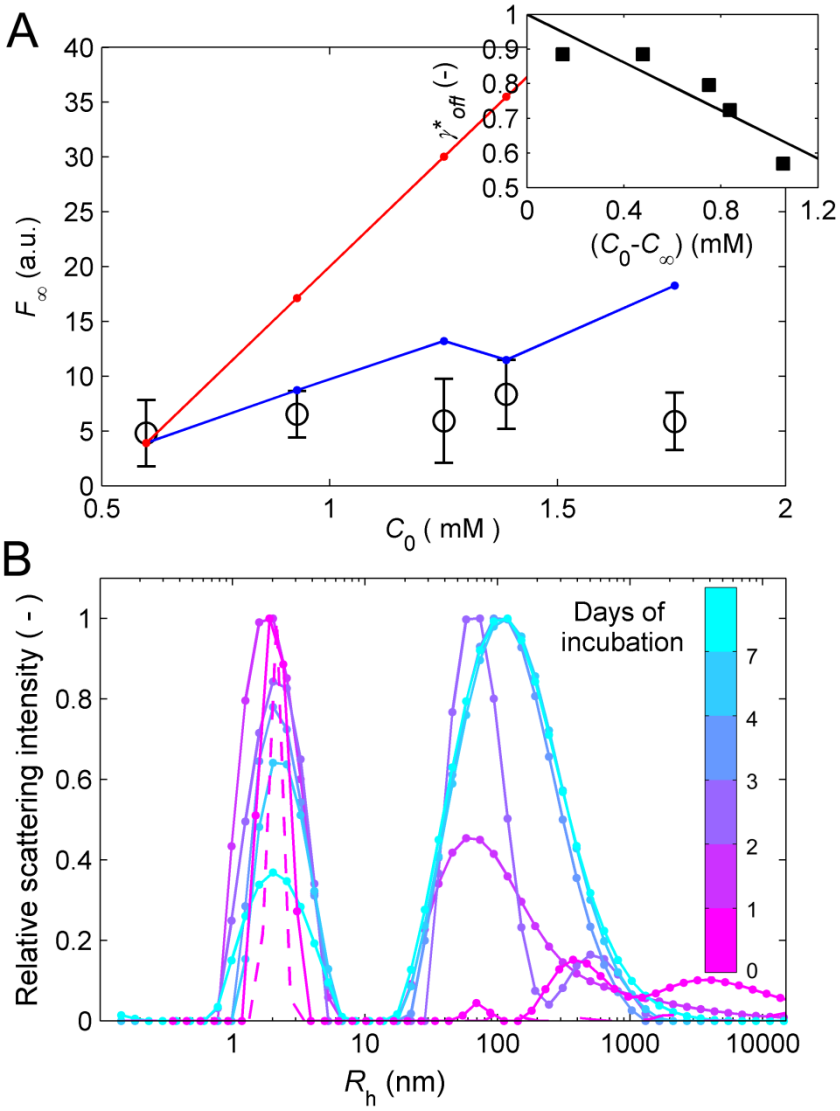


Figure 4.1. The effective concentration of soluble HEWL is influenced by the amount of insoluble HEWL aggregates in solution. (A) Open symbols and error bars: mean values and standard deviations of the end-point ThT fluorescence F_∞ measured for different initial HEWL concentrations C_0 at pH 1.6 and 60 ºC. Dots connected by red line: in the absence of solvation-in effects the end-point signal is expected to linearly increase with C_0 as the amount of amyloid fibrils is given by the (C_0- C_a*) difference and the amyloid solubility C_a* is constant. Dots connected by blue line: since C_a* increases with C_0 (Table 4.1) the theoretical F_∞ (C_0) dependency becomes attenuated. This variation alone is not sufficient to match the measured trend. Inset: solvation-in effects illustrated by decreasing activity coefficients γ_{off}* as the amount of total aggregates (C_0- C_∞) increases (symbols). The solid line is a linear fit to the data with slope −0.35 mM−1 and the y-intercept set to 1. (B) Distribution of hydrodynamic radii R_h obtained from DLS measurements after different periods of

incubation (color bar). Solid lines: incubation of 1.76 mM HEWL samples at pH 1.6 and 60 ºC[167]. Dashed line: freshly prepared 3.57 mM HEWL samples at pH 4.7. The first aggregates are formed upon acidification with a size distribution centered near 70 nm that is in good agreement with previous morphological analysis of the insoluble HEWL precipitates using advanced microscopy technique[167].

The DLS measurements in Figure 4.1B indicate that the first HEWL aggregates are formed upon acidification with a size distribution centered near 70 nm. This approximate size falls within the range of 20-150 nm observed for HEWL precipitates using Atomic Force Microscopy and Transmission Electron Microscopy[167]. The combination of these observations with the protein depletion data shown below (Figure 4.2) indicate that the formation of insoluble off-pathway aggregates starts immediately after incubation at pH 1.6 and 60 ºC, and dominates thereafter over any other process of aggregation.

Kinetic and thermodynamic effects caused by OPA:

To validate our description of amyloid fibril formation in the presence of OPA, the extended CLM was tested against HEWL aggregation progress curves measured by two distinct, yet complementary, spectroscopic methods: UV absorption and ThT fluorescence. The obtained results are represented in Figure 2, with each panel corresponding to different initial HEWL concentrations.

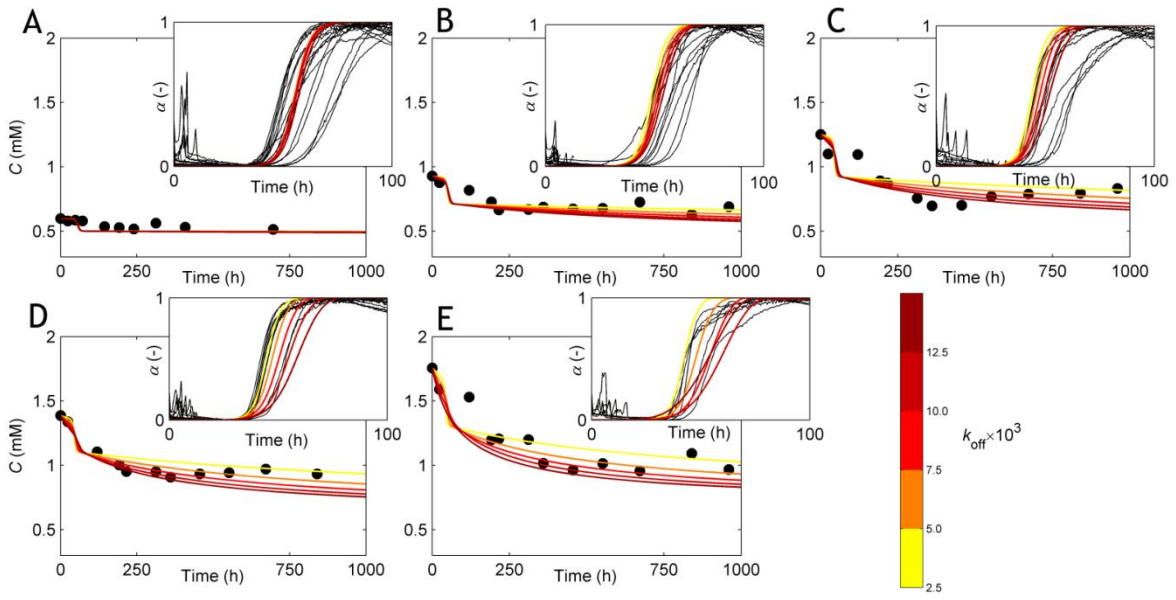


Figure 4.2. The CLM extended to consider OPA and solvation effects is able to predict protein depletion and amyloid fibrillization curves measured for different HEWL concentrations. Symbols: variation with time of the soluble HEWL concentration measured at pH 1.6 and 60 ºC for (A) C_0=0.60 mM, (B) 0.93 mM, (C) 1.25 mM, (D) 1.39 mM and (E) 1.76 mM HEWL. Colored lines: predictions of the CLM for corresponding values of C_0 using k_a =0.24 h^{-1}, k_b =1.0×10^{-7} and the values of k_{off} indicated by the color bar. Insets: normalized ThT fluorescence increase

measured (black lines) and simulated (different colored lines) under the same conditions as the *C(t)* curves. The variability of the ThT fluorescence progress curves can be described by changeable values of the nucleation rate constants k_{off} (color bar) and k_b (not shown).

It is remarkable that a single set of k_a, k_b, k_{off} fitted parameters is able to describe each pair of independent results in Figure 4.2, given that protein depletion (main graphs) and amyloid fibrillization (inset graphs) not only require different techniques to be followed as they take place over very distinct time scales. As documented by the $\alpha(t)$ progress curves, amyloid fibrils are principally formed within the period of 30–90 hours, whereas HEWL depletion generally starts before and continues long after this interval of time. Because amyloid fibrillization accelerates protein clearance from solution (see the initial inflections in the theoretical $C(t)$ curves) the two types of progress curves cannot be dissociated from each other. It is no less relevant that our simulation is able to capture the effect of changing HEWL concentration (Figure 4.2A to Figure 4.2E) considering the intricate role that C_0 has on the end-point fluorescence, duration of the lag phases and limit aggregation rates. On account of the stochastic nature of nucleation steps, the corresponding amyloid (k_b) and off-pathway (k_{off}) rate constants are adequately described by intervals of values as an alternative to single fits. By admitting the range of k_{off} values given in the color bar of Figure 4.2, the observed variability of experimental results can be particularly well described for higher HEWL concentrations (Figures 4.2D and 4.2E). Under these conditions, amyloid fibrils are a smaller fraction of the total aggregates and, therefore, OPA plays a more important role. The high variability of amyloid progress curves observed in the insets of Figures 4.2A to 4.2C can also be replicated by assuming an interval of k_b values around the fitted value (data not shown). The margin of error admissible for the CLM parameters is large enough to incorporate the influence of $\sigma_{a,0}$ and $\sigma_{off,0}$ on the concentration-independent rate constants $k_a/\sigma_{a,0}$, $k_b/\sigma_{a,0}$ and $k_{off}\sigma_{a,0}/\sigma_{off,0}$. According to the data in Table 1, $\sigma_{a,0}$ is comprised between 0.20 and 0.36 whereas $\sigma_{a,0}/\sigma_{off,0}$ is comprised between 0.17 and 0.59.

The fitting results are used in Figure 4.3 to describe the influence of the initial HEWL concentration on amyloid fibril formation. Amyloid conversion is expressed in

Figure 4.3A as the amount of fibrils effectively produced over the maximum amount expected in the absence of OPA (α_a).

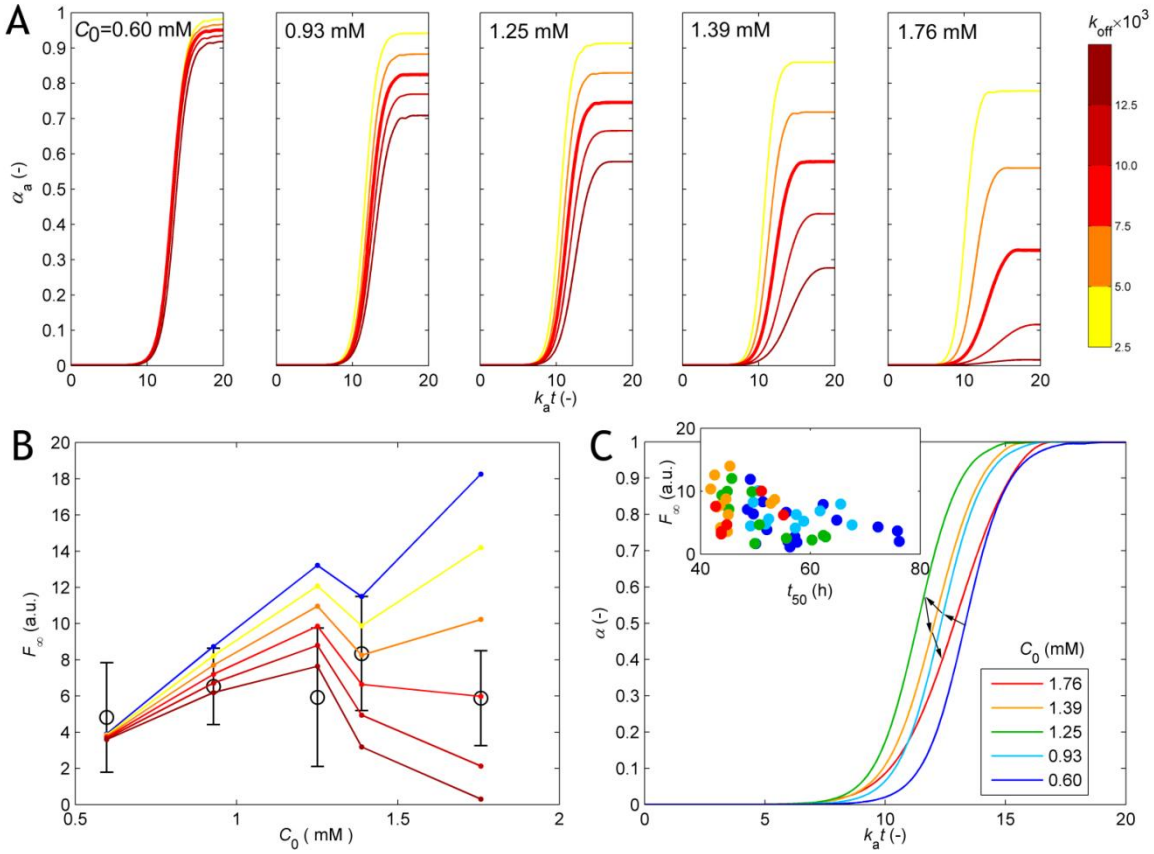


Figure 4.3. The presence of OPA and crowding-in effects attenuate the C_0-dependency of amyloid kinetic coordinates. (A) Fractional amyloid signal (α_a) represented as a function of the dimensionless time ($k_a t$) as predicted by the extended CLM using the same nucleation rate constants k_b and k_{off} as in Figure 2 and the values of C_0 indicated in each graph. (B) Open symbols and error bars: same experimental data as in Figure 4.1. Different colored dots connected by lines: predicted end-point amyloid signals (F_∞) obtained from the product of the maximum fluorescence values expected in the absence of OPA (blue dots) and the limit values of α_a calculated for long reaction times ($\alpha_{a,\infty}$) using the values of k_{off} indicated by the color bar. (C) Simulation data represented in terms of the normalized amyloid signal ($\alpha=\alpha_a/\alpha_{a,\infty}$) for C_0=0.60 mM (blue), 0.93 mM (light blue), 1.25 mM (green), 1.39 mM (orange) and 1.76 mM (red) after choosing a fixed value of k_{off} (7.5×10^{-3}). Arrows indicate increasing C_0 values. Changing the value of C_0 had no major effect on the duration of the lag phases or on the limit aggregation rates. Inset: Experimental data obtained during HEWL aggregation showing the variation of the endpoint ThT fluorescence emissions with the time required to reach 50 % completion (t_{50}) - same color code as in the rest of the panel.

The successive graphs in Figure 4.3A show that increasing protein concentrations favors the formation of off-pathway aggregates over amyloid fibrils. This kinetic promotion of OPA concurs with the thermodynamic solvation-in effect to complete the explanation of the weak variation of the end-point fluorescence with C_0 that was

previously identified in Figure 4.1A. The simulated end-point signals obtained in Figure 4.3A are shown in Figure 4.3B to reproduce very well the measured trend. Although a counterintuitive possibility, negative scaling laws are also admitted to occur when OPA is more prevalent - higher k_{off} values in Figure 4.3B. Additionally, the summary of fitted curves presented in Figure 4.3C suggests non-trivial behaviors of the half-life coordinates, with higher HEWL concentrations not necessarily meaning shorter aggregation times t_{50} or faster aggregation rates v_{50} . Formerly appointed as OPA signatures[167], the peculiar variations of t_{50} and v_{50} with C_0 are here shown to justify the simultaneous occurrence of solvation-in effects in order to explain the whole experimental data in a quantitative manner. The inset in Figure 4.3C addresses the variability of F_∞ as the result of the competition between amyloid and off-pathway formation[13,166,167]. As some of the soluble protein also produces ThT-negative precipitates, the stochastic properties of nucleation steps are propagated into poorly reproducible F_∞ values. Faster amyloid nucleation rates (i.e. higher k_b values) are expected to simultaneously provoke more intense end-point signals and shorter induction times. Only partially perceptible in the inset of Figure 4.3C, this inverse proportion is likely affected by changeable off-pathway nucleation rates that scatter the measured F_∞ values to a greater extent.

Supplementary validation experiments:

An immediate consequence arising from the extended CLM is that protein aggregation can be reactivated by decreasing the fraction of insoluble aggregates in solution. This hypothesis was tested during protein depletion experiments involving mixture volume expansion, which, as described in Material and Methods and graphically illustrated in Figure 4.4A, was carried out without changing the HEWL concentration, temperature or pH. We expected and confirmed that lower fractions of insoluble aggregates would reduce the solvation effects and promote additional HEWL aggregation. Although unaffected at the instant of the perturbation, the subsequent decrease of the molar HEWL concentration confirms that protein aggregation was boosted (Figures 4.4B to 4.4F). Similar tests involving ThT fluorescence monitoring evidenced no amyloid signal increase after the mixture volume expansion (data not shown) indicating that only OPA was promoted. As the experiments carried out for higher protein concentrations also produced more insoluble aggregates, the addition of clear HEWL solutions had, in those cases, more pronounced aggregation promoting effects (compare Figures 4.4B to 4.4F).

This is explained as the result of larger gains in the thermodynamic activity of HEWL when higher fractions of insoluble aggregates are diluted to one half; in the limit case of clear incubating solutions (without aggregates) no variation on the effective HEWL concentration would be induced and the aggregation propensity would remain unchanged. The same CLM parameters previously demonstrated to describe multiple kinetic results in Figures 4.2 and 4.3B are now shown to effectively predict the new HEWL depletion data represented in Figures 4.4B to 4.4F. Notice that the successful fittings would not be possible without extending the off-pathway-CLM to include activity coefficients seeing that the aggregation boost is caused by deliberately changing the magnitude of the salvation-in effects.

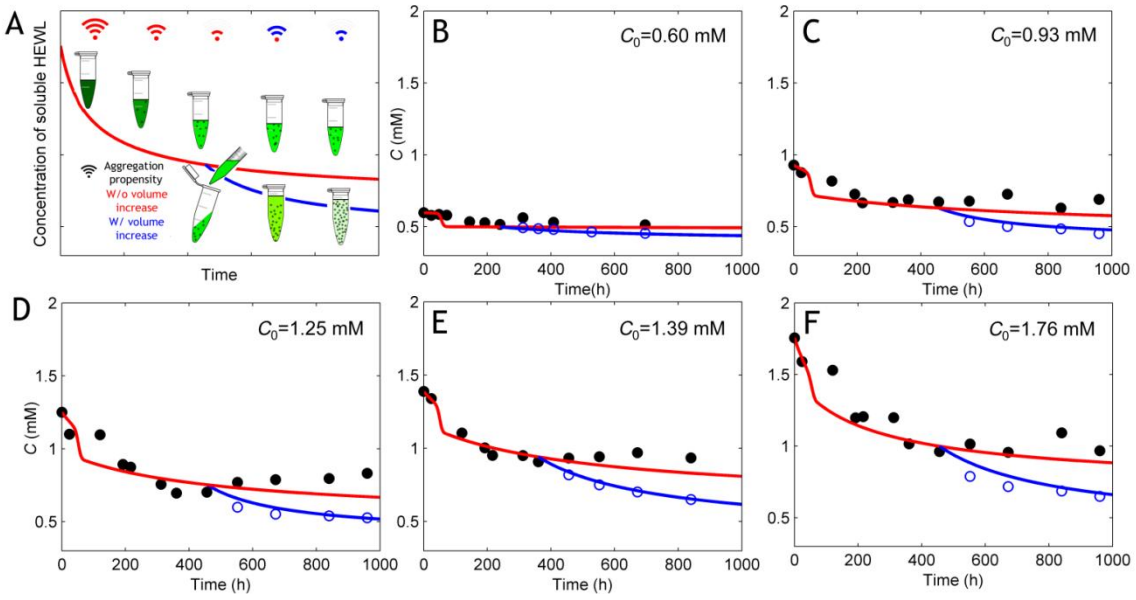


Figure 4.4. Protein aggregation is reactivated after decreasing the fraction of insoluble aggregates in solution. (A) Schematic variation the aggregation propensity (Wifi symbol) and soluble HEWL concentration (lines) with (blue) and without (red) expansion of the mixture volume. Plastic microtubes represent independent samples taken for analysis after different periods of incubation. At a fixed time instant, a number of samples were subject to volume expansion by the addition of 1.0 ml of clear protein solution prepared to the same HEWL concentration of the incubated mixture. (B-F) Lines: Protein depletion curves simulated using the same set of CLM parameters as in Figure 4.2 for the experiments started with (B) C_0=0.60 mM, (C) 0.93 mM, (D) 1.25 mM, (E) 1.39 mM and (F) 1.76 mM HEWL; same color code as in (A). A selection of fixed k_{off} values of (B) 12.5×10^{-3}, (C) 10.0×10^{-3}, (D) 12.5×10^{-3}, (E) 7.5×10^{-3} and (F) 7.5×10^{-3} was made from the color bar list in Figure 4.2. Filled circles: same experimental data as in Figure 4.2A to 4.2E, respectively. Open circles: variation of soluble HEWL concentration measured after the reaction volume expansion.

4.4 Discussion

We have shown that a set of three model parameters is enough to quantitatively explain multiple and surprising amyloid aggregation results in the presence of off-pathway crowding agents. The theoretical framework provided by CLM is confirmed using a repertoire of kinetic data measured during ThT fluorescence and protein depletion experiments of HEWL at low pH (1.6) and high temperature (60 C). By explaining the amyloid and OPA progress curves measured at different HEWL concentrations, the new model is also quantitatively describing puzzling scaling laws of half-life coordinates (both t_{50} and v_{50}) and end-point amyloid signal (F_∞) with C_0. Despite their singularity, similar behaviors have been identified by other authors and in other systems than lysozyme[142,144,166,174]. By successfully characterizing the changes on the relative importance of amyloid and off-pathway aggregation as different conditions are tested, the unexpected high variability of F_∞ is also understood as the result of stochastic variations on the nucleation rates; poorly reproducible protein aggregation is a recurring problem while studying different amyloidogenic proteins[108,145,175,176]. Given the different and independent experimental data analyzed, the number of degrees of freedom was reduced to a point where the CLM was actually used to predict (rather than to fit) the outcome of new experiments involving mixture volume expansion. These experiments confirmed that protein aggregation can be reactivated by decreasing the fraction of insoluble aggregates in solution. Crucial for the overall quality of the validation tests, the solvation-in effect introduces a thermodynamic dimension to the role of off-pathway aggregates during amyloid fibrillization. The common practice in past and recent literature of assuming the monomer concentration alone as the driving force for phase transition gave rise to a series of known inconsistencies between theoretical predictions and measured kinetic data[34,167]. By expressing the nucleation and growth steps as a function of supersaturation, the CLM is exempted from those limitations and is able to accommodate the effect of crowding agents on protein solubility. Figure 4.5A schematically presents a description of indirect solvation effects as responsible for altering the tendency of soluble protein to form aggregates according to the availability of water molecules in solution. Using as reference the enthalpic solvation equilibrium (top), the introduction of amyloid fibrils and off-pathway aggregates (right side) is a disturbance to this balance by recruiting protein molecules to the solid-liquid interface. As a result, the initially saturated polypeptide (intermediate color

shade) becomes undersaturated in solution (lighter color shade) while the surface-adsorbed macromolecules remain under phase equilibrium. Should the departing point consist of a supersaturated solution, the addition of aggregates would seed the formation and growth of more amyloid fibrils given that surface adsorption is a precursor step of integration. Macromolecular crowding, on the contrary, is known to upshift the effective concentration of the amyloidogenic protein and promote aggregation[151,169]. From the analogy with the salting-out effects provoked by electrolytes[177,178], the availability of water molecules is represented to decrease in the presence of macromolecular crowders (left side of Figure 4.5A). The formation of the hydration shell around the added molecules is responsible for the increased thermodynamic activity of the soluble protein (darker color shade). This 'solvation-out' mechanism is independent on whether the external agent is a polymer or its monomeric equivalent[159].

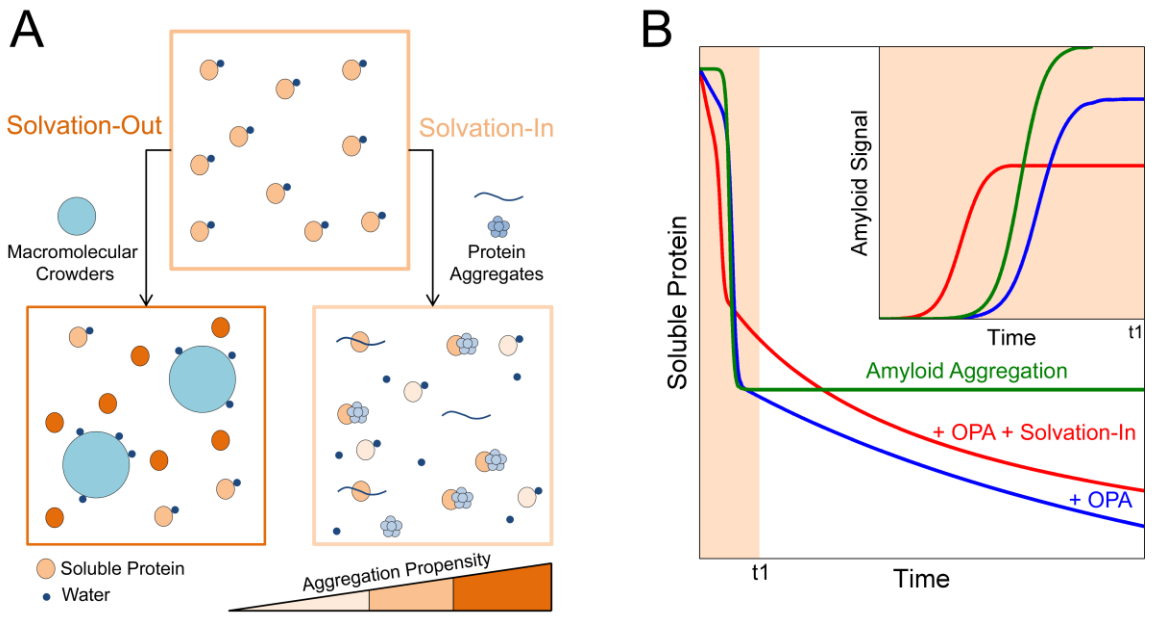


Figure 4.5. Indirect solvation effects during amyloid fibril formation. (A) Graphic illustration of solvation-out and solvation-in effects provoked by macromolecular crowding agents and protein aggregates, respectively. The propensity of protein molecules to aggregate is higher when the availability of water decreases (solvation-out) and vice-versa (solvation-in). (B) Illustrative protein depletion curves in the absence of OPA (green line), in the presence of OPA without solvation-in effects (blue line) and in the presence of OPA with solvation-in effects (red line). Inset: Corresponding variation of the amyloid signal with time. Shaded areas comprise an equal period of time (t1).

Although the simplified scheme does not include conformational and volume excluding effects resulting from molecular confinement, it provides a conciliatory view of macromolecular crowding not exclusively dependent on hard-sphere, entropic

interactions[159,179–181]. From the parallel with the present work, we speculate that macromolecular crowding can equally be characterized (i) using effective supersaturation levels as the driving force for protein aggregation, and (ii) measuring the soluble protein concentrations after long incubation times to estimate thermodynamic activity coefficients. The kinetic outcomes expected in those cases do not necessarily oppose the general tendency summarized in Figure 4.5B for off-pathway aggregates, which, differently from macromolecular crowders, are not considered to induce volume excluding effects. Added to this, OPA is kinetically competing with amyloid fibrillization for soluble protein, whereas classic crowding agents are generally inert molecules. The competition between amorphous and amyloid aggregation events explains part of the differences between green (without OPA) and blue (with OPA) progress curves represented in Figure 4.5B, namely the earlier concentration drop in the presence of OPA and the lower end-point signal shown in the inset. The protein depletion curve in blue also represents an increased reaction extent caused by the formation of off-pathway aggregates which, at least for HEWL, are less soluble than amyloid fibrils. The existence of solvation-in effects does not change the interfacial activity of the insoluble phase (not represented in Figure 4.5B) but increases the apparent concentration of protein under equilibrium conditions (red progress curve in Figure 4.5B). As a result, although lower amounts of total aggregates are produced in the presence of solvation-in effects, the final thermodynamic availability of protein molecules remains unaltered in the two OPA cases shown in the graph. Besides affecting the magnitude and variability of the end-point amyloid signal, OPA interferes with the fibrillization kinetics in a complex manner that is not totally evident in the inset of Figure 4.5B. At the same time that the rates of amyloid nucleation and growth are slowed, the rate of protein depletion is increased (due to the extra phase transition step) - while the former outcome delays the onset of the final plateau phase, the latter anticipates its arrival. Solvation-in effects are, in turn, responsible for faster decrease of the supersaturation levels leading to faster overall amyloid kinetics but lower amounts of fibrillar aggregates. As the previous analysis of Figure 4.3 demonstrated, both OPA and solvation-in effects contribute for weakening the influence of the initial protein concentration on typical kinetic coordinates such as the duration of the lag-phase, limit aggregation rate and end-point amyloid signal.

In addition to the experimental observables here summarized, non-linear $\alpha/(1-\alpha)$ vs. time log-linear plots complete the list of signatures associated to OPA. We believe that

a new class of off-pathway regulators can hereafter be screened based on their ability to amplify/attenuate these kinetic indicators[167]. The finding that off-pathway species additionally deplete soluble protein through solvation-in mechanisms is expected to have a twofold importance for the development of new drugs targeting protein aggregation: the total aggregates content emerges as an important variable affecting the scalability of in-vitro results, and OPA might promote (unpleasantly) surprising losses of physiological function. First, the recurrent principle that crowded environments promote aggregation may not apply when the crowding agent consists of proteinaceous aggregates. It remains to be ascertained how differently packed milieus interfere with the efficacy of drug candidates. Second, therapeutic strategies deflecting the aggregation pathway to non-amyloidogenic precipitates might also promote loss-of-function cascades involving more extensive aggregation, lower amounts of soluble protein, and protein inactivation. On the other hand, OPA may arguably represent a competing pathway with protein folding and a defense mechanism of the cell in response to deregulated protein homeostasis[182,183]. All these factors have to be balanced against possible toxicity gains resulting from reduced amounts of amyloid fibrils.

4.5 Conclusions

Originally derived based on generic phase transition principles of supersaturation-driven nucleation and growth, the CLM is now ready to address OPA as an important peculiarity of amyloid fibril formation determining toxicological aspects of protein aggregation. This theoretical framework predicts that OPA directly reduces the formation of amyloid fibrils, a consequence that might be considered good or bad depending on which type of aggregates is more toxic. Off-pathway amorphous precipitates showing lower solubility than structured fibrils also contribute for lowering the availability of the amyloidogenic protein in biological media. The OPA mechanism of action involves kinetic and thermodynamic aspects that were thoroughly confirmed using protein depletion and ThT fluorescence data measured during amyloid fibril formation of HEWL. Solvation-in effects, here introduced to explain the shortage of soluble protein in the presence of off-pathway aggregates, also account for variations on the apparent solubility values as different initial HEWL concentrations were considered and decisively contribute to quantitatively validate kinetic signatures of OPA such as the characteristic scaling laws of

the half-life coordinates. Directly influenced by solvation-in effects, the end-point amyloid signal is highly variable in the presence of OPA and is weakly influenced by the initial protein concentration. More than providing a good numerical adjustment to measured data, the extended CLM accounting OPA and indirect solvation effects is confirmed as a predictive model that needs a reduced number of parameters to anticipate multiple experimental outcomes. Illustrating this, the reactivation of protein aggregation observed after expanding the mixture volume was successfully simulated using a set of three model parameters previously fitted to independent kinetic data. Compatible with high-throughput screening methods of drug discovery, the new findings are expected to close the gap between in vitro and in vivo results and contribute for the design of new therapeutic strategies involving OPA.

Chapter 5

Application example:

Screening of on/off-pathway inhibitors of aSyn aggregation

Summary

While the pathogenic role of aSyn amyloid fibrils is not yet clear, evidences suggest that intermediate species may be responsible for toxicity in the aSyn aggregation process. Extensive research has recently been focused in developing new inhibitors producing a direct or indirect effect on oligomeric species. As a result, most aggregation/toxicity inhibitors under study work by accelerating aggregation, removing or remodeling toxic oligomers or reducing fibril fragmentation, but it is not yet understood at which stages of oligomerization they are effective. On the other hand and remaining to be targeted, off-pathway aggregates (OPA) suggest being important regulatory/toxic agents in aSyn fibrillization. A new knowledge-based strategy by using the CLM as a tool for the characterization of regulatory and anti-aggregation/toxicity mechanisms is presented in the following work. Preliminary tests of benzothiazole derivatives screening were developed in order to illustrate the application of this new method of analyses. Benzothiazole-2-carbonyl chloride suggested affecting the formation of OPA. Further studies will be required in order to present any conclusion about its effect on aSyn fibrillization.

5.1 Introduction

Alpha synuclein (aSyn) is a 140-amino acid protein encode by the SNCA gene and generally expressed at higth levels in neurons, in central and peripheral nervous system[184]. aSyn is an intrinsically disordered protein, presenting instability in its secondary and tertiary structure[116,185,186]. Evidences have recently suggested that aSyn in vivo may co-exist in equilibrium as a partially folded monomer with a variety of other dynamic oligomeric species. Although showing stability when adopting a helically folded tetramer structure, as a disordered monomer and dynamic dimer aSyn tends to aggregate[187].

The aggregation of aSyn into amyloid fibrils, filamentous assemblies of proteins in which the polypeptide backbone is arranged in a characteristic cross-beta-sheet structure[121], is related with a diverse group of neurodegenerative disorders known as synucleinopathies in which Parkinson disease (PD) is included[188]. PD is a chronic neurodegenerative disease characterized by the loss of dopaminergic innervation through death of neurons affecting one in 100 people over the age of 60[189]. In PD, deposited aSyn amyloid fibrils are the main component of Lewy bodies of the substantia nigra, constituting the second most common nerve cell pathology, after the neurofibrillary lesions of Alzheimer's disease[186,190].

Extensive research has been invested in recent years for the development of new therapeutic strategies for PD and other synucleinopathies. Small molecule compounds are considered as one of the available possibilities to inhibit protein assembly into insoluble amyloid aggregates and intermediated species[185]. These compounds could inhibit fibril aggregation by stabilizing the monomers, stabilizing the non-propagating aSyn dimmers, interfering with the aggregation path preventing the appearance of toxic oligomers, inhibiting aSyn oligomerization, preventing interactions between monomers and protofibrils (fibril growth inhibition), disassembling preformed fibrils and/or preventing amyloid-membrane interactions[31]. Although numerous aggregation inhibitors are currently being tested under different preclinical and clinical phases, the mechanism of action of this type of therapy is not completely understood. A renewed approach to drug discovery is needed favoring the 'in-cerebro' component and creative innovation over blind technology[191,192]. A knowledge-based strategy and its application are presented in this work by using the CLM as a tool for the characterization of anti-aggregation mechanisms. The central idea is to use basic principles of the physics of phase transitions on the quest of

new therapeutic options based on the inhibition of supramolecular assembly. High-throughputscreens performed during the lead discovery and optimization phases are often reported to identify candidate aggregation inhibitors that are later dismissed during the expensive preclinical and clinical phases of the drug development pipeline. We ascribe much of this productivity shortfall to the phenomenon of apparent inhibition, i.e., non-specific thermodynamic effects that reduce the amount of aggregates produced in-vitro but not in-vivo. This is because cells and tissues media are in generalmuch more buffered against small variations in composition than the solutions prepared in lab. Changing the thermodynamic activity by the addition of salts is a simple example of such interactions[130]. Hits having specific aggregation-rate lowering effects are promising lead compounds and deserve further investment. Impressive as they may seem, results showing decreased amounts of produced aggregates should be handled with care especially if they result from indiscriminate thermodynamic effects. The CLM will be able to discriminate between true and apparent inhibitors with the goal of identifying false positives earlier in the drug development pipeline[130]. Basic science is therefore used during the lead discovery phases in order to improve the much-lamented poor success rates on preclinical and clinical phases.

However, references from PD suggest hallmark inclusions may be a protective mechanism developed by neurons to prevent the accumulation of the pathogenic species, oligomers[185,193].

5.2 Experimental Procedures

5.2.1. Chemicals

pT7-7 vector encoding for wild type (wt) α-synuclein (aSyn) was a gift from Hilal Lashuel (Addgene plasmid # 36046). ThT was obtained from Sigma-Aldrich (St Louis, MO, USA) and used as received. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was obtained from AppliChem (Darmstadt, Germany). Other materials and reagents were analytical grade and obtained from Sigma-Aldrich (St. Louis, MO, USA) and Merk KGaA (Darmstadt, Germany).

5.2.2. Plasmid amplification, DNA purification and glycerol stocks

Plasmid pT7-7 aSyn wt was received as transformed bacteria in stab culture format (*E.coli* DH5α) and stored at 4ºC. Using a sterile toothpick, bacteria were streaked on a Luria broth (LB) agar plate containing 100 µg/ml ampicillin to grow single colonies. An overnight 50 ml LB liquid culture inoculated from a bacteria single colony (100 µg/ml ampicillin) was growth.

A plasmid DNA isolation miniprep kit (NucleoSpin Plasmid, Macherey-Nagel GmbH & Co.KG., Düren, Germany) was used for plasmid DNA isolation from the bacterial cells. DNA was quantified spectrophotometrically at 260 nm. The purity of the sample was assessed by analyzing the A260/280 and the A260/230 ratios.

Glycerol stocks were prepared since bacteria on an LB agar plate can be stored at 4°C for no longer than a few weeks.500 µl of the overnight culture were added to 500µl of 50% glycerol solution in 2 mL screw top tubes. Screw top tubes were gently mix, frozen in liquid nitrogen and stored at -80 ºC.

5.2.3. Transformation of BL21 (DE3) *E.coli*with Plasmid DNA

BL21 (DE3) *E.coli*competent cells were obtained from Labclinics (Barcelona, Spain) and stored at -80 ºC. Transformation of BL21 (DE3) *E. coli*cells with plasmid DNA was developed using the heat shock method.

50 ng of DNA were added to 50 µl of competent cells. After 30 minutes incubation on ice, the mixture of chemically competent bacteria and DNA was exposed at 42°C for 45 s (heat shock) and then placed back on ice. Right after the addiction of 1 ml of LB medium, transformed cells were incubated at 37°C for 1 hour under agitation. A centrifugation at 13.000 rpm (room temperature) for 2 minutes was carried out and 950 µl of the supernatant were discarded. 100 µl of transformed bacteria were plated on a LB agar plate containing 100 µg/ml ampicillin and incubated overnight at 37 °C.

5.2.4. Recombinant wt aSyn production

BL21 (DE3) *E. coli* cells were transformed as explained before. An overnight 10 ml LB liquid culture inoculated from a bacteria single colony (100 µg/ml ampicillin) was growth and used to inoculate 1 l of LB medium containing 100 µg/ml ampicillin. The 1 l culture was incubated at 37°C with shaking (180 rpm) until the OD_{600} reached 0.6. Recombinant protein expression was induced by the addiction of 0.5 mM IPTG and incubated in the same conditions (37°C, shaking 180 rpm) for 4h. Bacterial cells were harvested by centrifugation for 20 minutes at 10,000 rpm and 4 ºC (Sorvall LYNX 6000

Superspeed Centrifuge and Fiberlite F12-6x500 LEX Rotor, Thermo Scientific, MA, USA). After removal of the supernatant, the cell pellets were frozen at -80 ºC in the centrifugation bottles (Nalgene, Thermo Scientific, MA, USA).

5.2.5. Purification of wt aSyn

Non chromatographic purification steps

Under tap water, cell pellets were pushed with a spatula and transferred to a clean flask kept on ice. They were then resuspended in 100 ml lysis buffer (10 mM Tris·HCl pH 8; 1 mM EDTA). The sample was divided into 10 ml volumes, kept on ice and ultrasonicated (Digital Sonifier Branson SLPe; 4C15, 40KHz converter) with an output power of 150 watts, 60 % amplitude, applied in 30 s pulses followed by a 30 s pause for a total ultrasonication time of 6 min. Cell debris were then pelleted by centrifugation at 17,000 rpm in a Sorvall RC-5B centrifuge (Sorvall SS-34 Rotor, Thermo Scientific, MA, USA) for 20 min at 4 ºC. An ammonium sulfate precipitation step was performed by adding 0.166 g/ml ammonium sulfate and stirring of the solution during 1 hour at 4 ºC for sample homogenization. Supernatant was kept after centrifugation in the same condition as described before (17,000 rpm, 20 min, 4 ºC). A second ammonium precipitation step was performed by adding 0.129 g/ml ammonium sulfate, to selectively precipitate the aSyn protein. Sample was stirred for 1 hour at 4 ºC followed by centrifugation (17,000 rpm, 20 min, 4 ºC). Supernatant was discarded and the pellet resuspended in 80-100 ml phosphate buffered saline (PBS). After 20 minutes of boiling at 90 ºC to denature and precipitate most cellular proteins, another centrifugation was carried out (17,000 rpm, 20 min, 4 ºC). Supernatant (containing aSyn) was kept at 4 ºC, ready for chromatographic purification steps.

Chromatographic purification steps

The supernantant obtained from the non chromatographic purification was loaded at 5 ml/min onto a HiTrap Q HP anion exchange 5 ml column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden), equilibrated with 50 mM phosphate buffer pH 7.4. ASyn was eluted around 300 mM NaCl (at 5 ml/min) by applying a linear gradient, in 50 mM phosphate buffer pH 7.4, over 15 column volumes. The elution fractions were analyzed by SDS-12% PAGE, and fractions containing aSyn (~ 17 kDa bands) were pooled and concentrated to 2 ml volume using 10 kDa Amicon Ultra-15 centrifugal filters (Merk

Millipore, Darmstadt, Germany). ASyn sample was further purified by gel filtration chromatography using a HiLoad 16/600 Superdex 75 prep grade column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) equilibrated with 50 mM phosphate buffer pH 7.4, 200 mM NaCl. ASyn was eluted at 1 ml/min. All column chromatography was performed on an ÄKTA prime plus chromatography system (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). Pure aSyn fractions were combined and concentrated to 800-1000 µl volume using 10 kDa Amicon Ultra-4 centrifugal filters (Merk Millipore, Darmstadt, Germany).Protein concentration was determined by measuring the absorbance at 275 nm using an extinction coefficient of 5,600 M^{-1} cm^{-1}. Aliquoted protein was stored at -80 ºC after freezing in liquid nitrogen.

5.2.6. Gel Electrophoresis (SDS-PAGE) and Immunoblotting

ASyn samples were diluted in loading buffer, boiled for 5 min at 95 ºC and separated on 12% SDS Mini Protean TGX stain-freeprecast gels (Bio-Rad, CA, USA).Gels were run at 40 mA for 45min in an electrophoresis tray and visualized in a ChemiDoc XRS+ System (Bio-Rad, CA, USA). For blot analysis, the proteins were transfered onto nitrocellulose membrane (Merk Millipore, Darmstadt, Germany) using a tank blot system (Bio-Rad, CA, USA) and blocked overnight with 5% w/v non fat dry milk in PBS-T (0.01% (v/v) Tween 20 in PBS) at 4 ºC. The membrane was probed with the primary antibody (rabbit polyclonal antibody aSyn (C-20)-R obtained from Santa Cruz Biotechnology, CA, USA) at a dilution of 1:1000 for 1 h at room temperature. After 4 PBST washes, the membrane was incubated with the secondary antibody (polyclonal goat anti-rabbit immunoglobulins/HRP obtained from Dako, Glostrup, Denmark) at a dilution of 1:5000, for 1 h at room temperature and protected from light. The immunoblot was finally washed 4 times with PBS-T and 4 times with PBS, incubated with Luminata Crescendo Western HRP Substrate (Merk Millipore, Darmstadt, Germany) and visualized in a ChemiDoc XRS+ System (Bio-Rad, CA, USA).

5.2.7. Aggregation assays: ThT Fluorescence

ThT fluorescence kinetic measurements were carry out at 37 ºC in 96-well microtiter microplates (Thermo Scientific, MA, USA) in a CHAMELEON ™ V Microplate Reader (Hidex Co., Turku, Finland) at an excitation wavelength of 440 nm and an emission wavelength of 485 nm. ThT stock solution was prepared by dissolving the dried powder in 50 mM phosphate buffer pH 7.4 and filtered through a sterile 0.45 µm pore size PES membrane filter (Jet biofil, Seoul, Korea). The concentration was

determined by absorbance measurements at 411 nm using an extinction coefficient of 22000 $M^{-1}cm^{-1}$. Samples of 120 µl, containing a 3 mm diameter glass bead per well (Glaswarenfabrik Karl Hecht, Sondheim, Germany), with a final ThT concentration of 70 µM and aSyn concentration between 50 and 200 µM were sealed with 100 µl of paraffin oil. Measurements were recorded every 900 s, after sample homogenization by 300s shaking. Data were background-corrected for the ThT fluorescence of the respective solvent in the absence of protein.

5.2.8. Screening of small molecule compounds: ThT Fluorescence

ThT derivative compounds and polyphenols were first solubilized in dimethylsulfoxide (DMSO) to stock concentrations of 1 mM and kept at -20 ºC.

Samples of 120 µl were prepared as described before for aggregation assays using a fixed protein final concentration of 140 µM and 70µM ThT. 10µl of freshly prepared compound solutions (obtained by dissolving the 1 mM stocks in 50 mM phosphate buffer pH 7.4) were added to the samples to final concentrations of 5 to 200 µM. DMSO final concentration did not exceed 1% (v/v). Samples were sealed with 100 µl of paraffin oil and measurements were recorded every 900 s, after sample homogenization by 300s shaking.

5.3 Results and discussion

5.3.1 Protein production

Recombinant aSyn was produced in BL21 (DE3) *E. coli* cells and purified by a extensively used and high-yield protocol including non-chromatographic and chromatographic steps (see Experimental section).

Non-chromatographic steps are based on sample boiling to remove non-aSyn proteins by precipitation. Since aSyn is an unfolded protein, sample boiling does not affect its structure, remaining in solution, while natively folded cytoplasmic proteins thermally denature and precipitate.

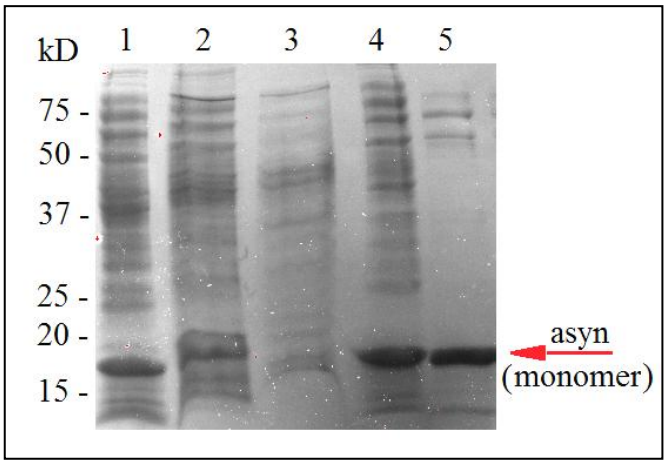


Figure 5.1. SDS-PAGE analysis of the aSyn non chromatographic purification steps. Lane 1, supernantant remaining after cell lysis. Lane 2, supernatant after the first ammonium sulfate precipitation step. Lane 3, supernatant after the second ammonium sulfate precipitation step (aSyn is precipitated in this step). Lane 4, pellet after the second ammonium sulfate precipitationstep resuspended in PBS. Lane 5, supernatant after boiling.

Figure 5.1 shows the SDS-PAGE analysis for the different steps of the non-chromatographic purification process. The predominance of themonomeric state was observed. Although aSyn has a molecular mass of 14.5 kDa, it migrates on SDS-PAGE with an apparent mass of ~ 17 kDa as a consequence of the poor binding of its N-terminal acidic tail to SDS molecules[187] (Figure 5.1).

Chromatographic steps consisted in an ion exchange and size exclusion chromatography (Figure 5.2 and 5.4, respectively). The obtained yield of purified protein was about 10 mg/l of *E. coli* culture.

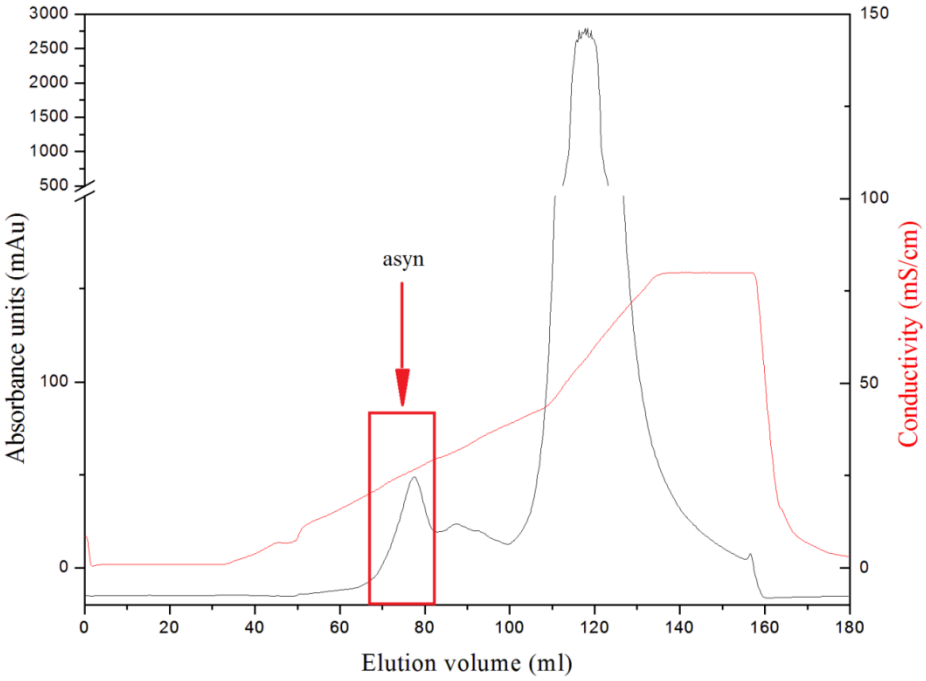


Figure 5.2. Purification of aSyn. Anion exchange chromatography of the supernatant after boiling on a 5 ml HiTrap Q HP column. Absorbance is represented in black and conductivity in red. ASyn peak was eluted around 300 mM NaCl, comprising fractions from an elution volume of 68 ml until 82 ml.

After the anion exchange chromatography, the purity of theaSyn peak's fractions was further confirmed by SDS-PAGE (Figure 5.3) before continuing with the size exclusion chromatography (Figure 5.4). ASyn was eluted showing a sharp peak comprising from 53 to 64 ml. Fractions were also further analyzed by SDS-PAGE (Figure 5.5), showing bands of ~ 17 kDa.

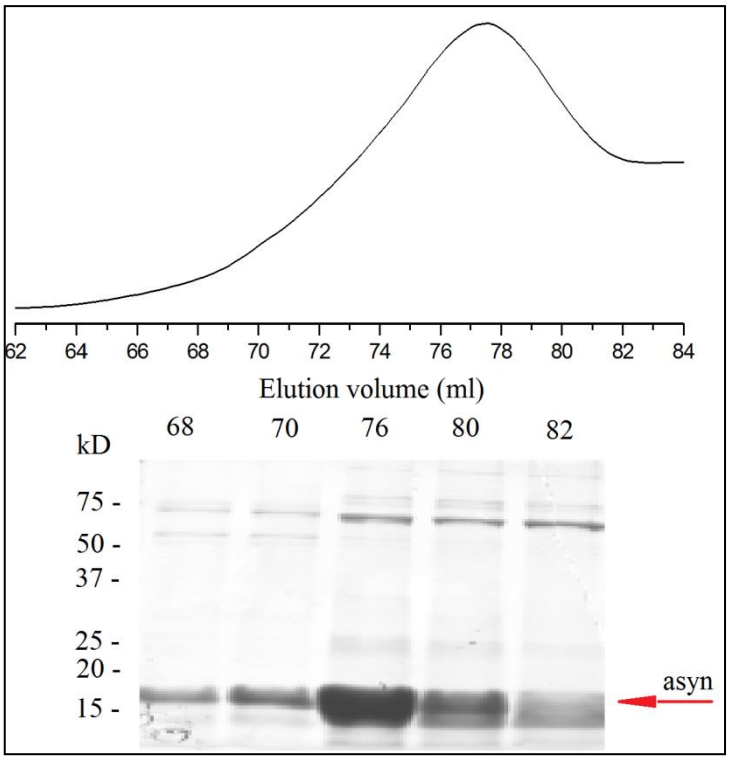


Figure 5.3. SDS-PAGE analysis of the fractions comprised in the aSyn peak (anion exchange chromatography, Figure 5.2). Bands of ~ 17 kDa were observed.

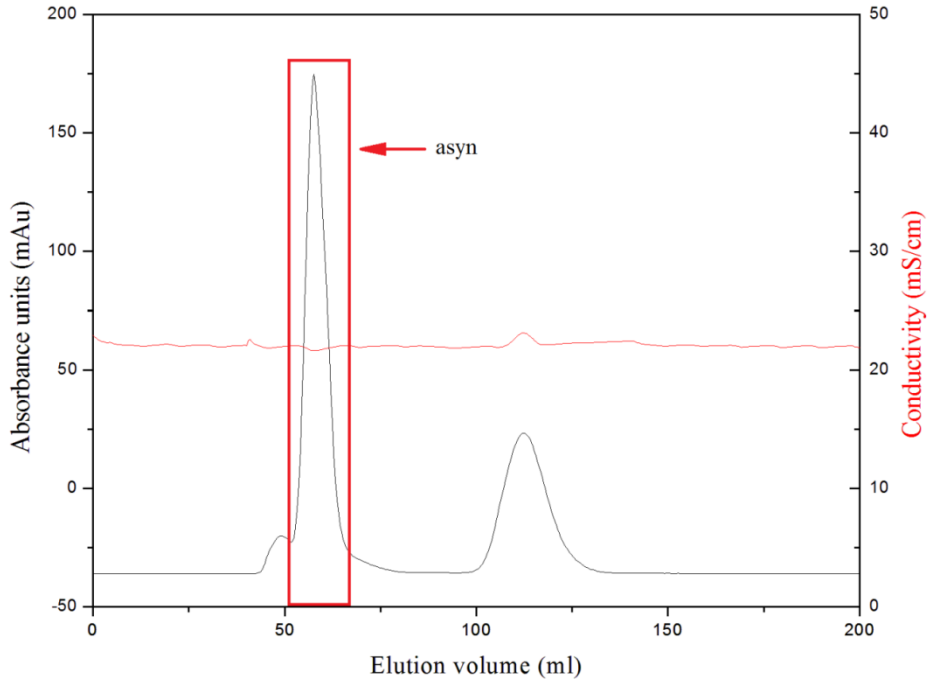


Figure 5.4. Purification of aSyn. Size exclusion chromatography of aSyn purified by previous anion exchange chromatography. Absorbance is represented in black and conductivity in red. The sharp peak comprising from 53 to 64 ml elution volume, corresponds to monomeric aSyn.

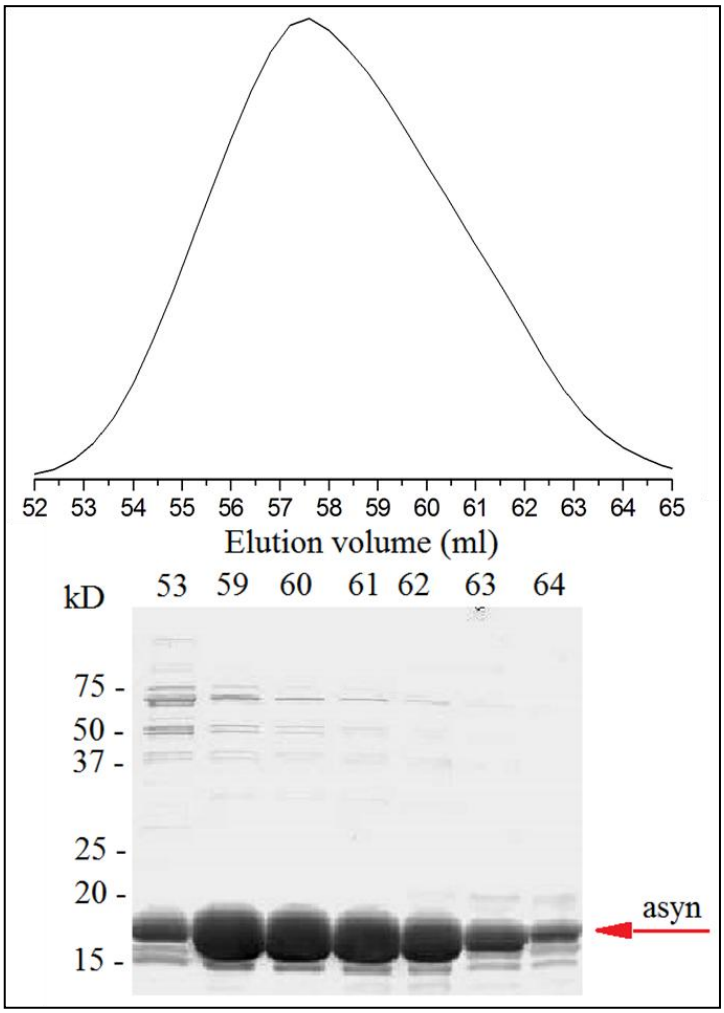


Figure 5.5. SDS-PAGE analysis of the fractions comprised in the monomeric aSyn peak (size exclusion chromatography, Figure 5.4). Bands of ~ 17 kDa were observed.

5.3.2 ASyn aggregation kinetics

ASyn aggregation is very sensitive to solution conditions such as temperature, pH, salt concentration and/or solution cofactors. The kinetics of aSyn aggregation were measured for 5 different aSyn initial concentrationsat pH 7.4 and 37 ºC. Experimental data were represented as the normalized intensity of ThT fluorescence with time (Figure 5.6A) and the concentration-dependent half-life coordinates (Figure 5.6B and C). Glass beads were used to increase agitation, promoting hydrophobic-water interfaces[116].

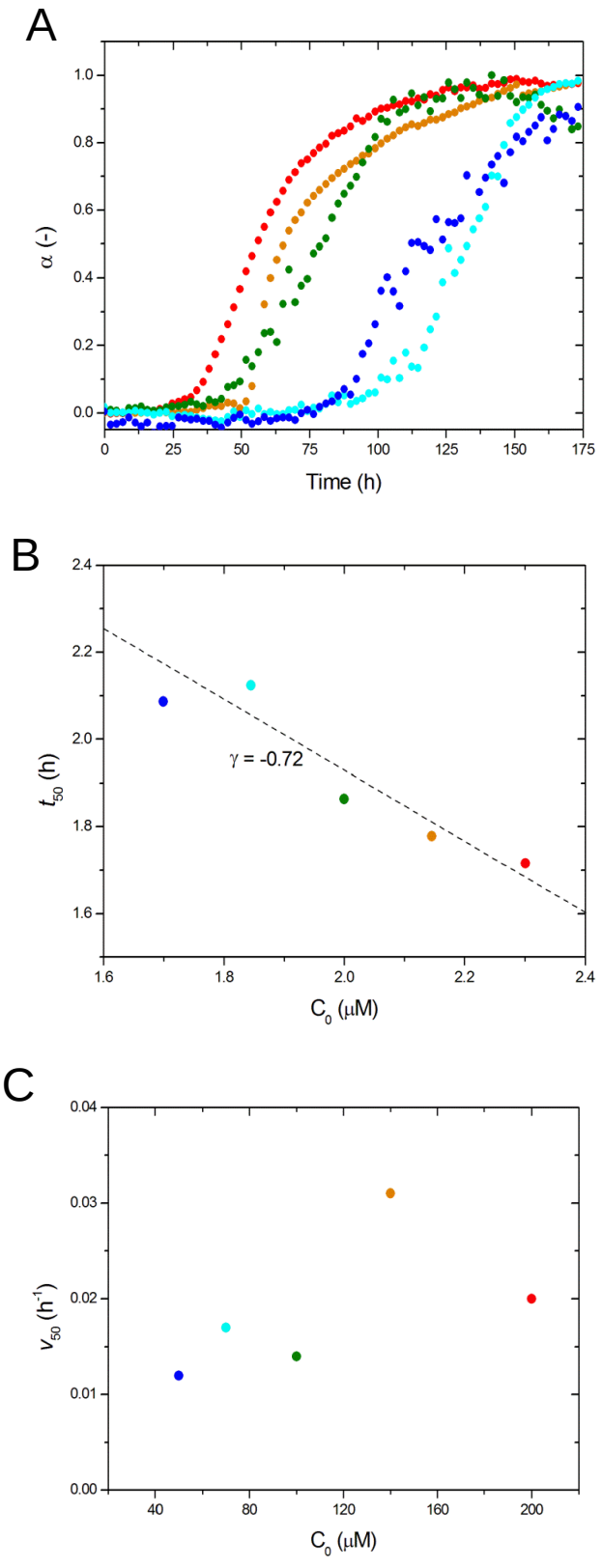


Figure 5.6. Influence of the initial protein concentration on the kinetics of amyloid fibril formation of aSyn at pH 7.4 and 37 ºC. (A) ThT fluorescence increase during the aggregation of 50 μM (blue), 70 μM (light blue), 100 μM (green), 140 μM (orange) and 200 μM (red) aSyn represented in normalized units as a function of the incubation time. (B, C) Influence of the initial aSyn concentration (B) on the time

required to reach 50% completion (t_{50}), log-log scale, and (C) on the aggregation rate at the same instant (v_{50}), linear-linear scale.(B) The linear fit (dashed line) indicates an exponential scaling factor γ of −0.72.

The weak concentration dependence showed by the lag-time and the half-life coordinates, as well as the obtained absolute value for the scaling factor lower than 1, suggest the parallel formation of intermediate and off-pathway species. Obtained data of ThT fluorescence was represented as a function of the incubation time in order to illustrate the fluorescence endpoint signals, F_∞ (Figure 5.7).

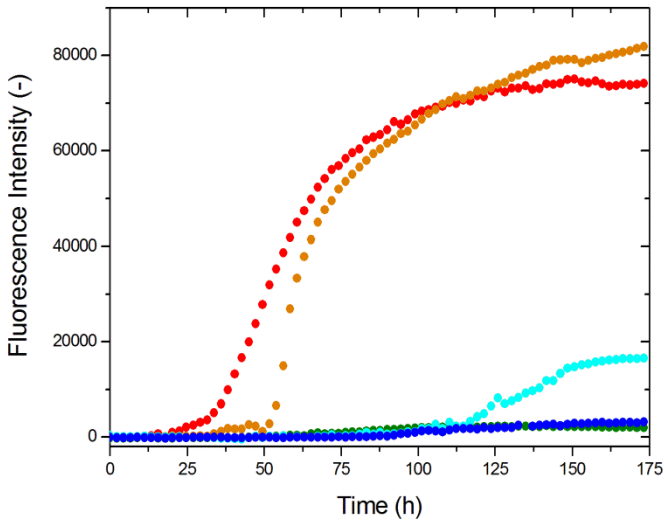


Figure 5.7. ThT fluorescence increase during the aggregation of 50 µM (blue), 70 µM (light blue), 100 µM (green), 140 µM (orange) and 200 µM (red) aSyn represented as a function of the incubation time.

The variability of the F_∞ showed in Figure 5.7 is another evidence of the formation of intermediates and off-pathway species during aSyn amyloid fibrillization. For a complete qualitative analysis, log-linear representations of the measured $\alpha/(1-\alpha)$ with the incubation time are presented (Figure 5.8).

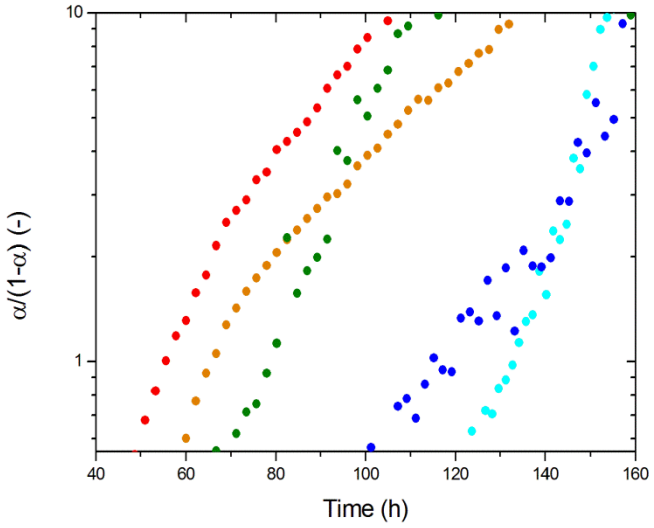


Figure 5.8. The measured aggregation curves of aSyn in Figure 5.7 are represented in the modified coordinates (same color code as in Figure 5.7). The concave phase observed for higher asyn initial concentrations suggest the existence of OPA.

Linear trends starting after a short period of time until the end of the reaction will be representative of conventional hyperbolic/sigmoidal amyloid aggregation curves. The nonlinear concave trends during the fast growth stages are considered as atypical kinetic signatures, suggesting the presence of off-pathway aggregates. Additional scaling studies are necessary in order to consolidate this hypothesis.

5.3.3 Screening of small molecule compounds

A selection of benzothiazole derivatives was done based in the specific binding mode of ThT to amyloid fibrils (Table 5.1). Their chemical structure is showed in Figure 5.9. Amyloid fibril aggregation kinetics were measured in the presence of these compounds for 140 μM aSyn at pH 7.4 and 37 ºC.

The polyphenols epigallocatechin-3-gallate (EGCG) and curcumin were also assayed as reference inhibitors since their effect by interfering directly with amyloid fibril formation from several peptides and proteins had been recognize (Figure 5.9).

Table 5.1. List of benzothiazole derivative compounds assayed as possible inhibitors.

Compound Nº	**Compound Name**
1	Benzothiazole
2	Benzothiazole-2-carboxaldehyde
3	Benzothiazole-5-carboxylic acid
4	Ethyl benzothiazole-2-carboxylate
5	2-(2-Aminophenyl)benzothiazole
6	2-(4-Aminophenyl)benzothiazole
7	2-(4-Piperidinyl)-1,3-benzothiazole
8	2-Hydroxy-1,3-benzothiazole-6-carboxylic acid
9	2-Methyl-1,3-benzothiazole-6-carboxylic acid
10	4,5,6,7-Tetrahydro-1,3-benzothiazole-2-carboxylic acid
11	Benzothiazole-2-carbonyl chloride
12	2-(Chloromethyl)benzothiazole
13	2-(Bromomethyl)-1,3-benzothiazole
14	2-Chloro-1,3-benzothiazole-6-carboxylic acid
15	6-(Hydroxymethyl)benzothiazole
16	2-(1*H*-Pyrazol-3-yl)-1,3-benzothiazole
17	2-(1*H*-Pyrrol-1-yl)-1,3-benzothiazole-6-carboxylic acid
18	2-pyrrolidin-2-yl-benzothiazole
19	PGL-135 hydrochloride monohydrate

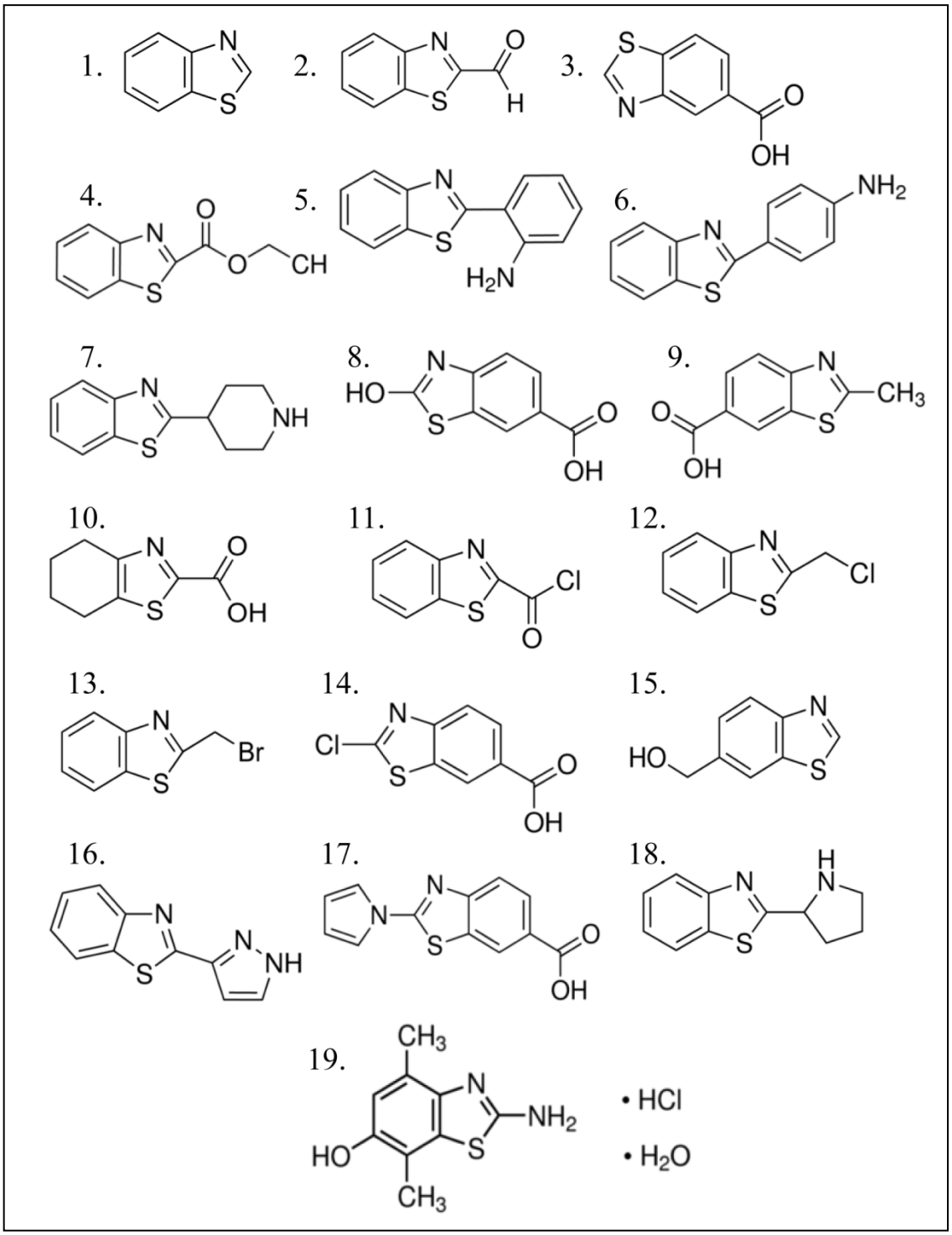


Figure 5.9. Core structure of the benzothiazole derivatives assayed as possible inhibitors (names related to the numbers in Table 5.1).

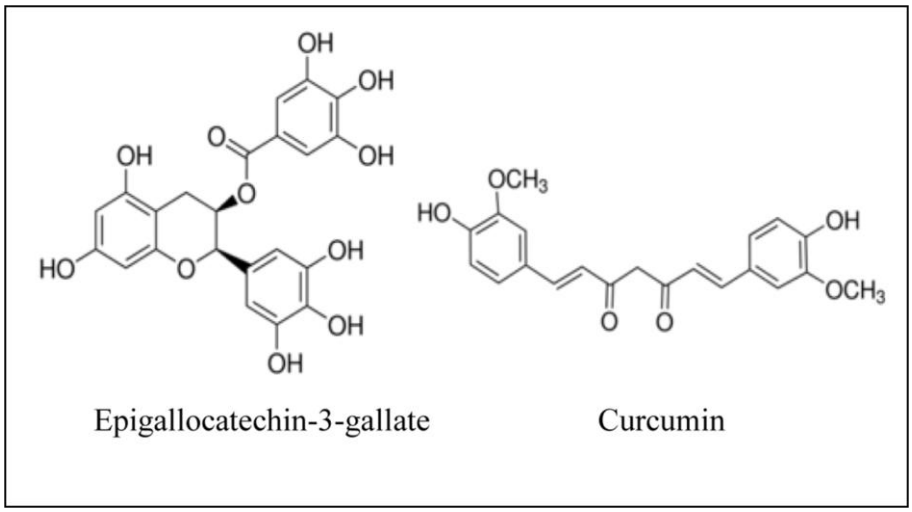


Figure 5.10. Core structure of polyphenols, EGCG and curcumin, assayed as reference inhibitors during aSyn amyloid fibrillization.

On-pathway inhibition

The impact that external factors, as small molecule compounds, can produce on amyloid fibrillization is attributed by the previously described CLM (Chapter 2) to kinetic or thermodynamic reasons. In the first case, alterations in the duration of the initial lag time and/or the maximum aggregation rate are observed. A specific molecular binding between the inhibitor and the folded protein is required to affect the rate constants for nucleation and/or growth. In the second case, the end plateau level of aggregation is affected since alterations are produced in the effective protein concentration of the system, although the kinetic constants will remain unchained. No specific chemical affinity is required for disturbing the system equilibrium and show marked effects on amyloid fibril formation. We call the compounds that only show a thermodynamic effect in *in vitro* experiments "apparent inhibitors". Those compounds will be discarded as amyloid inhibitors in later phases of the drug development process where cellular proteostasis will regulate the non-specific disturbances. The CLM offers a simple and practical possibility to recognize "apparent" or "true" inhibitors against experimental data. By fitting the CLM to experimental data through eq. 2.17 (Chapter 2), the growth and nucleation-to-growth rate constants are quantified.

An example of this analysis is illustrated below for the cases of 2-(2-aminophenyl) benzothiazole (Figure 5.11) and EGCG (Figure 5.12). Other assayed compounds analysis is shown in Supplementary Information Chapter 5 (Appendix).

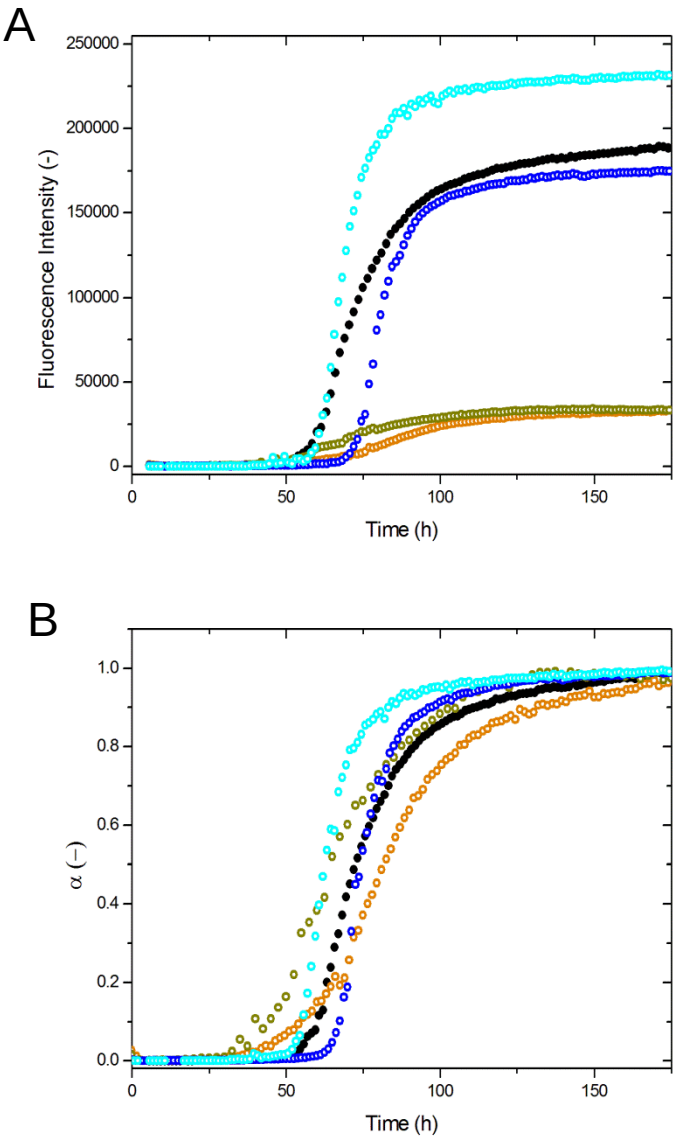


Figure 5.11. Influence of 2-(2-aminophenyl)benzothiazole on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (A) ThT fluorescence increase represented as a function of the incubation time. (B) ThT fluorescence increase represented in normalized units as a function of the incubation time.

The end plateau level of aSyn aggregation seems to be affected by 2-(2-aminophenyl)benzothiazole (Figure 5.11A). In order to analyze if any variation occurred in the kinetic rate constants, the obtained fibril mass increase was plotted in normalized units as a function of time (Figure 5.11B). No kinetic effects appear to be produced by 2-(2-aminophenyl)benzothiazole, indicating a case of apparent inhibition.

In this preliminary test of screening of inhibitors, the assayed ThT derivative compounds yielded no suggestions of true inhibition on aSyn fibrillation. Contrarily, some cases of apparent inhibition were easily detected, following the approach illustrated as an example in Figure 5.11.

When kinetic effects are observed in the representation of the normalized fibril mass increase, by fitting the CLM to thedata through eq 2.17 thekinetic rate constants will be quantified.As an example of this kind of approach, the CLM was recently used on the development and validation of fluorinated peptides that prevent the aggregation of amyloid-beta (aβ) peptide in Alzheimer's disease[194].

If fibrillization is blocked by the compound, the measured aggregation curves will show a linear trend instead sigmoidal profiles showingan initial lag phasefollowed by the subsequent exponential growth phase. EGCG and curcumin have dose-dependently shown to produce this effect on aSyn fibrillization[31,60]. Those compounds may bind to oligomeric species of aSyn, blocking interactions between aSyn and aSyn toxic oligomeric species[31].

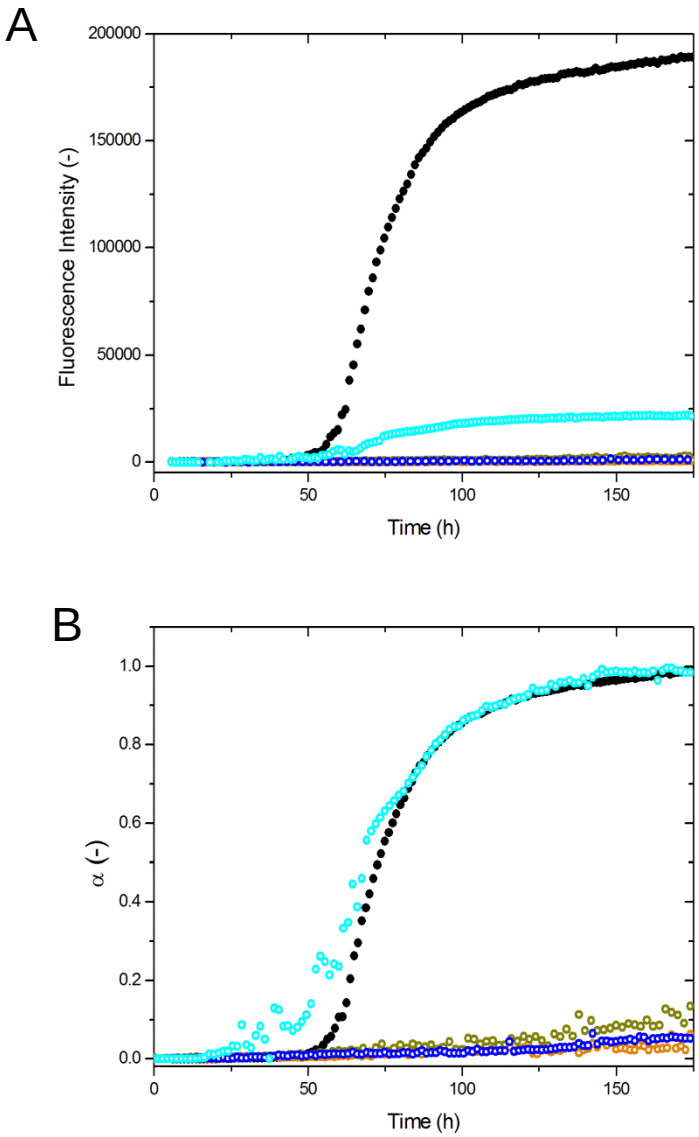


Figure 5.12. Influence of EGCG on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (A) ThT fluorescence increase represented as a function of the incubation time. (B) ThT fluorescence increase represented in normalized units as a function of the incubation time.

Our results appear to confirm what had been previously shown in the literature[31,126]. Figure 5.12 illustrates the influence of different EGCG concentrations on the kinetics of aSyn amyloid fibril formation. Linear trends were obtained in the presence of the highest EGCG concentrations assayed. The lowest concentration did not appear to be enough to inhibit aSyn aggregation. These results allow us to classify EGCG as a true inhibitor. Compounds showing true inhibition effects (producing alterations in the kinetic rate

constants or completely blocking the process) will be selected for further screens. An increase of the success rate on the search for new inhibitory compounds will be observed by applying this knowledge-based strategy in high-throughput screens.

Off-pathway regulation

aSyn fibrillization is a process involving the formation of many different intermediate species and off-pathway aggregates[193]. Evidences suggest that part of the intermediate aggregates may be the pathogenic species, turning them one of the main targets of the inhibitory strategies in which extensive research has been invested[195,196]. EGCG and baicalein are examples of compounds producing the inhibition of aSyn aggregation by acting directly on the toxic oligomers[126,197]. As a consequence of this effect, an increase in the formation of non-toxic OPA may be produced. Special care will be required with compounds that block or reduce aSyn fibrillization not acting directly on the toxic aggregates since an increase in toxicity may be possible[195]. Other compounds showing dose-dependent inhibitory effects, as curcumin, seem to increase the solubility of aSyn monomers and affect as oligomerization as fibrillization, even its mechanism of action is not yet clear[31]. Alterations in OPA during the screening of inhibitors of aSyn fibrillization may give crucial information about the kind of inhibition and the mechanism of action produced by the compounds. OPA seem to act as regulatory/toxic agents in those specific cases of inhibition. We have applied our new CLM-tools in this preliminary test of benzothiazole derivatives screening in order to identify the influence that the compounds may have in the formation of OPA during aSyn fibrillization.

By representing the measured $\alpha/(1-\alpha)$ as a function of time in log-linear scale, the kinetic profilesobtained in the presence of benzothiazole derivatives were analyzed. Possible effects of alterations in the formation of OPA may be detected by a direct comparison with the profile obtained for aSyn in the absence of inhibitors.

Results did not suggest an increase in the formation of OPA in any of the cases (Appendix, Supplementary Information Chapter 5). Concave profiles obtained from modified coordinates representation were in the same order of curvature as the one showed by aSyn in the absence of inhibitors. Contrarily, we found a case suggesting linear trends. In Figure 5.13, the influence of benzothiazole-2-carbonyl chloride on the kinetics of aSyn amyloid fibril formation and its representation in modified coordinates are presented.

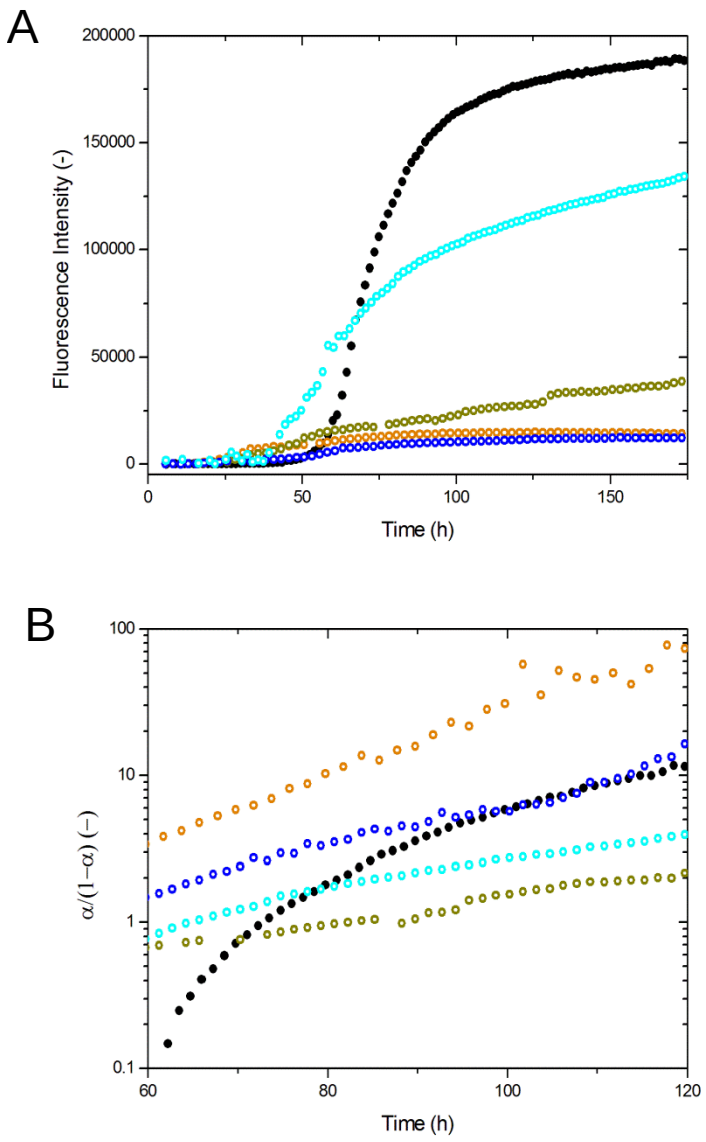


Figure 5.13. Influence of benzothiazole-2-carbonyl chloride on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5µM (light blue), 50µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (A) ThT fluorescence increase represented as a function of the incubation time. (B) Representation in modified coordinates, $\alpha/(1-\alpha)$ as a function of time in log-linear scale.

Linear trend observed in Figure 5.13B may suggest an effect of benzothiazole-2-carbonyl chloride on intermediate and/or off-pathway species. More detailed and specific studies are required in order to analyze what kind of aggregates are present, their toxicity, and to understand the possible mechanism of action of benzothiazole-2-carbonyl chloride.

The previous analysis illustrates an example showing the simplicity and easy applicability presented by our new proposed method to analyze the formation of OPA. This new tools may help for a better understanding of the mechanism involved in the interaction between small molecule compounds and protein aggregates during screening of inhibitors, contributing as well for the clarification of the mechanism comprised in protein fibrillization.

5.4 Conclusions

With this preliminary screening of compounds we show in a simple way how to analyze any possible effect produced on on-pathway and off-pathway aggregation during amyloid fibrillization. When focusing on on-pathway aggregation, to distinguish between thermodynamic or kinetic effects of external factors during amyloid fibril formation will be crucial. Chemical compounds having a thermodynamic effect induce different reaction affecting the effective protein concentration of the system. Those effects are not likely to work *in vivo* as they do *in vitro*, being this the reason why they are considered "apparent". Changing the solution activity of the protein by the addition of salts is a simple example of such effects. "True" inhibitors, on the contrary, have some kind of specific activity by retarding the kinetics of nucleation and/or growth of amyloid fibrillization. The CLM provides the tools to discriminate between apparent and true inhibitors in the early phases of the drug development process, but also, in its extended form (by the addiction of the off-pathway kinetic constant, Chapter 3), parallel nucleation steps can be characterize. We have also applied our new CLM-tools to detect OPA and the influence that benzothiazole derivatives compounds can have on it in this preliminary screening of compounds during aSyn fibrillization. Their potential as off-pathway modulators is shown by their effect attenuating or intensifying the OPA kinetic signatures. It provides a better understanding of the mechanism involved in the interaction between the compounds and protein aggregates as well as in the protein fibrillization process. This original approach for detecting and quantifying OPA modulators, as well as the ability to discriminate between "true and apparent" inhibitors, makes from the CLM a simple and easily applicable tool compatible with higth-throughput screening methods for drug discovery.

Chapter 6

General discussion and Future perspectives

General discussion and future perspectives

Supersaturation, key factor to drive the production of high quality protein crystals was previously and deeply studied by us as the driving force of this process, inspiring us in our decision of exploring common aspects between crystallization and amyloid fibrillization. In the meanwhile, an excellent paper from the group of Naiki and Goto was published at the end of 2011 showing the effects of ultrasounds on the promotion of nucleation and growth of amyloid fibrils[198]. The experimental approach followed by these authors was also very similar to the one that we were proposing and that was originally put forward by us in 2010 to crystallize proteins[5]. After this fact, and in parallel with our work, extensive research has been developed in the area. Not a coincidence, but a way to conclude the work comprised in this thesis should be consider as a high quality contribution to the advancement of knowledge in the *in vitro* amyloid aggregation research.

We have used well-established concepts such as the Classical Nucleation Theory and size-dependent growth to describe amyloid fibril formation according to the "Crystallization-Like Model" (CLM), proposed by us in 2012. We found that the CLM is able to describe adequately all types of amyloid fibrillization kinetics identified in our literature survey, which included hundreds of different time-course aggregation curves. It is notable that the solid biophysical basis of the CLM does not compromise its final simplicity so that it can be understood by non-scientific audience. The model parameters can be simply determined from the usually available values of the time required to reach 50% completion (t_{50}) and of the aggregation rate at that instant (v_{50}). While integrating the combined influence of seeding, nucleation, growth and fibril breakage, we are enthusiastic about this model which is very close to the generalized underlying principle governing the kinetics of amyloid fibrillization.

As a consequence of the increasing evidences of parallel processes taking place during amyloid aggregation and their apparent relevance for the development of amyloidosis and neurodegenerative diseases, techniques and strategies to detect them in a simple way become essential. Since 2015, the CLM was prepared by the addition of a third kinetic constant, to quantitatively estimate the amount of off-pathway aggregates using no

additional means than the traditional amyloid-specific markers. The numerical solutions of the three-paramenter CLM unveil the kinetic signatures characteristic of OPA. OPA mechanism of action showed to involve kinetic and thermodynamic aspects, producing solvation-in effects that account for variations on the apparent solubility of the system under study. Those new findings are significantly improving the gap between *in vitro* and *in vivo* studies driving the effort to a more successful result. The CLM is compatible and ready to apply in hit-to-lead screenings of drug candidates through the in-vitro kinetic assay. We show as an application example our preliminary test, a starting list of ~20 benzothiazole derivatives, candidate inhibitors of pathogenic aSyn aggregation.

The CLM may be shortly extended to describe size distribution offering more detailed information about oligomerization, the different stages of the process and species involved, making possible to determine pathological species and the desired properties and mechanism of action of aggregation inhibitors: the goal for "the amyloid inhibition approach".

As a result, basic science is hereby used to develop new therapeutics for amyloidosis and neurodegenerative diseases. We all agree that new therapeutic strategies may result from a better understanding of the biophysics behind neurodegenerative disorders. This is particularly valuable for the investigation of diseases such as Alzheimer's, Parkinson's, Prion or other protein-related amyloidosis. We expect that the CLM may be routinely used to characterize quantitatively and qualitatively new targets for amyloid disease therapeutics working as a complement to other structural and molecular studies.

Discussão geral e perspectivas futuras

Sobressaturação, factor essencial para impulsionar a produção de cristais de proteínas de alta qualidade foi estudado por nós anteriormente e em profundidade como a força diretriz deste processo, inspirando-nos na nossa decisão de explorar aspectos comuns entre cristalização e agregação amilóide. Enquanto isso, um excelente trabalho do grupo de Naiki e Goto foi publicado no final de 2011, mostrando os efeitos de ultrasons sobre a promoção da nucleação e crescimento de fibras de amilóide. A abordagem experimental seguida por estes autores foi muito semelhante ao que nos estavamos a propor e que foi originalmente apresentado por nós em 2010 para a cristalização proteinas . Após este facto, e em paralelo com o nosso trabalho, extensa investigação tem sido desenvolvida nesta área. Não é uma coincidência, mas sim uma maneira de concluir que o trabalho desta tese deve ser considerado como uma contribuição de alta qualidade para o avanço do conhecimento na investigação *in vitro* da agregação amilóide.

Neste trabalho usamos conceitos bem estabelecidos, tais como a Teoria Classica da Nucleação e o crescimento dependente do tamanho para descrever a formação de fibras de amilóide de acordo com o "Crystalization-Like Model" (CLM), proposto por nós em 2012. Nós descobrimos que o CLM é capaz de descrever adequadamente todos os tipos de cinéticas de formação de amilóide identificados na nossa bibliografia, incluíndo centenas de diferentes curvas de agregação recolhidas ao longo do tempo . É notável que a sólida base biofísica do CLM não compromete a sua simplicidade final, de modo que pode ser compreendido por público não científico. Os parámetros do modelo podem ser simplesmente determinados a partir dos valores normalmente disponíveis do tempo necessário para atingir 50 % de conclusão (t_{50}) e da taxa de agregação nesse instante (v_{50}). Combinando a influência de *seeding*, nucleação, crescimento e fragmentação, deixa-nos entusiasmados por ser muito próximo do princípio fundamental geral que rege a cinética de formação de fibras de amilóide.

Como consequência das crescentes evidências de processos paralelos que ocorrem durante a agregação amilóide e a sua aparente relevância para o desenvolvimento de amiloidoses e doenças neurodegenerativas, técnicas e estratégias para detectá-los de uma forma simples tornam-se essenciais. Desde 2015, o CLM foi preparado acrescentando uma terceira constante cinética, para estimar quantitativamente a quantidade de agregados *off-pathway* usando nao mais do que os tradicionais marcadores específicos de amilóide. As

soluções numéricas do CLM com três parametros desvendaram as assinaturas características das cinética de agregação de OPA. Os mecanismos de ação de OPA mostraram envolver aspectos cinéticos e termodinâmicos, produzindo efeitos de *solvation-in*, que alteram a solubilidade aparente do sistema em estudo. Estas novas descobertas estão a melhorar significativamente a diferença entre os estudos *in vitro* e *in vivo* fazendo que o esforço seja conduzido a um resultado mais bem sucedido. O CLM é compatível e está pronto para ser aplicado na descoberta *hit-to-lead* de novo fármacos através de ensaios cinéticos de aggregação realizados *in-vitro*. Neste trabalho mostramos como exemplo de aplicação um teste preliminar de *screening* de uma lista inicial de ~20 derivados de benzotiazol candidatos a inibidores da agregação patogénica da aSyn.

O CLM será brevemente desenvolvido para descrever a distribuição de tamanhos oferecendo informação mais detalhada sobre a oligomerização, as diferentes fases do processo e das espécies envolvidas, tornando possível determinar espécies patológicas e as propriedades desejadas e mecanismos de ação dos inibidores da agregação de amilóide: a meta da "abordagem para a inibição amilóide".

Como resultado, a ciência básica é usada neste trabalho para desenvolver novos agentes terapêuticos para doenças neurodegenerativas e amiloidoses. É consensual que novas estratégias terapêuticas poderão resultar de uma melhor compreensão biofísica das doenças neurodegenerativas. Isto é particularmente valioso para a investigação de doenças como Alzheimer, Parkinson, Prion ou outras amiloidoses envolvendo a agregação de proteínas. Espera-se que o CLM possa contribuit para a caracterizaçã quantitativa e qualitativa de novos alvos terapêuticos em doenças de amilóide, trabalhando como um complemento a outros estudos estruturais e moleculares.

Discusión general e perspectivas futuras

La sobresaturación es un factor clave en la producción de cristales de proteína de alta calidad y estudiada previamente y en profundidad por nosotros como la fuerza motriz de este proceso, nos inspiró en nuestra decisión de explorar los aspectos comunes entre cristalización y formación de fibras de amilóide. Entre tanto, una excelente publicación del grupo de Naiki y Goto apareció a finales de 2011, mostrando los efectos de los ultrasonidos para promover la nucleación y el crecimiento de fibras deamilóide. El enfoque experimental seguido por estos autores fue muy similar al que nosotros estábamos desarrollando y que fue originalmente propuesto por nosotros en 2010 en la cristalización de proteínas. Después de este hecho, y en paralelo con los avances de nuestro trabajo, extensa investigación en esta área de estudio ha sido desarrollada. No debe ser visto como una coincidencia, sino como una manera de concluir que el trabajo comprendido en esta tesis supone una importante contribución para el avance del conocimiento en la investigación *in vitro* de la agregación amilóide.

Hemos utilizado conceptos bien establecidos, como la Teoría Clásica de la Nucleación y el crecimiento dependiente del tamaño para describir la formación de fibras de amilóide de acuerdo con el "Crystallization-Like Model" (CLM), propuesto por nosotros en 2012. El CLM demostró ser capaz de describir adecuadamente todos los tipos de cinéticas deformación de fibras de amilóide identificados en nuestro estudio bibliográfico, incluyendo cientos de diferentes curvas de agregación recogidas a lo largo del tiempo. Es notable que la sólida base biofísica del CLM no comprometa su simplicidad final y pueda ser asequible incluso para la audiencia no científica. Los parámetros del modelo pueden ser simplemente determinados a partir de los valores generalmente disponibles del tiempo requerido para alcanzar el 50 % de la reacción (t_{50}) y la velocidad de agregación en ese instante (v_{50}). Combinando la influencia de “seeding”, nucleación, crecimiento y fragmentación, nos entusiasma su cercanía al principio fundamental general que rige la cinética de formación de fibras de amilóide.

Como consecuencia a las evidencias de la existencia de procesos de agregación paralelos que tienen lugar durante la agregación amilóide y su aparente relevancia en el desarrollo de las amiloidosis y enfermedades neurodegenerativas, el desarrollo de técnicas y estrategias que permitan detectarlos de manera sencilla se ha convertido en un objetivo esencial. Desde 2015, modificaciones en el CLM fueron realizadas mediante la adición de

una tercera constante cinética, permitiendo estimar cuantitativamente la cantidad de agregados off-pathwaymediante los métodos tradicionales de uso de marcadores de amilóide específicos. Las soluciones numéricas del CLM de tres parametros revelan la firma cinética característica de OPA. El mecanismo de acción de OPA mostró involucrar aspectos cinéticos y termodinámicos, produciendo efectos de exclusión molecular que afectan la solubilidad aparente del sistema en estudio. Estos nuevos hallazgos mejorarán significativamente el abismo entre los estudios *in vitro* e *in vivo*, conduciendo el esfuerzo a un resultado más exitoso. El CLM es compatible y está listo para ser aplicado en "hit-to-lead screenings" de candidatos a fármacos a través del seguimiento de cinéticas de agregación *in vitro*. Mostramos como un ejemplo de aplicación nuestro estudio preliminar, una lista inicial de ~20 derivados del benzotiazol, como candidatos a inhibidores de la agregación de aSyn patógena.

El CLM será próximamente extendido para comprender distribución de tamaños, ofreciendo de este modo información más detallada acerca de las diferentes especies involucradas en el proceso, así como de sus diferentes etapas. Esto hará posible determinar las especies patológicas y las propiedades deseadas y el mecanismo de acción de los inhibidores de la agregación: el principal objetivo del "panorama amilóide".

Como resultado, la ciencia básica es utilizada en este trabajo para el desarrollo de nuevas herramientas terapéuticas para las amiloidosis y enfermedades neurodegenerativas. Todos estamos de acuerdo en que las nuevas estratégias terapéuticas pueden ser el resultado de una mejor comprensión de los procesos biofísicos detrás de las enfermedades neurodegenerativas. Esto es particularmente valioso para la investigación de enfermedades como Alzheimer, Parkinson, Prion u otras amiloidosis relacionadas con la agregación de proteínas. Esperamos que el CLM se utilice de forma rutinaria para caracterizar cuantitativa y cualitativamente nuevos objetivos para terapias de enfermedades de amilóide, trabando como complemento a otros estudios estructurales y moleculares.

Appendix

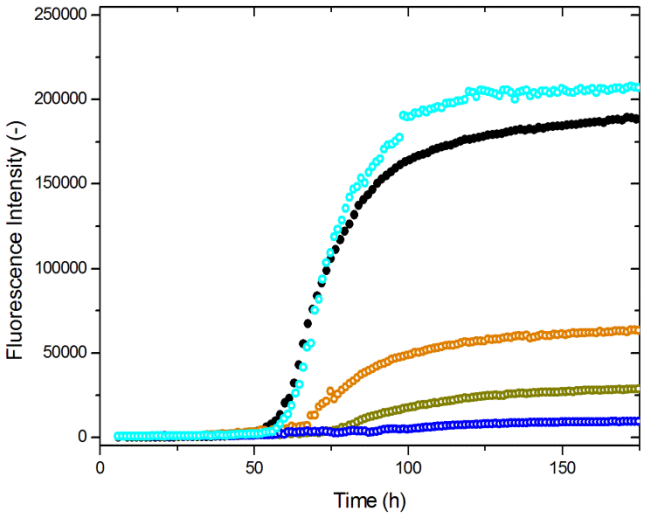
Supplementary Information

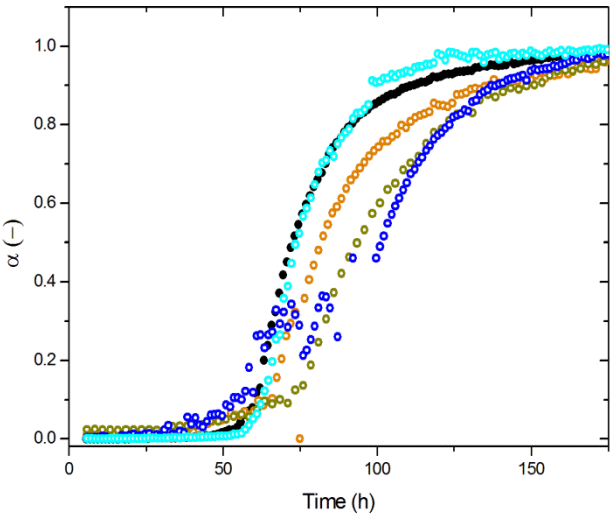
Chapter 5

Screening of on/off-pathway inhibition of aSyn aggregation

Amyloid fibril aggregation kinetics (140 μM aSyn at pH 7.4 and 37 ºC) were measured in the presence of a selection of benzothiazole derivatives and the polyphenols epigallocatechin-3-gallate (EGCG) and curcumin. The obtained data and corresponding analysis are shown bellow.

1. Benzothiazole





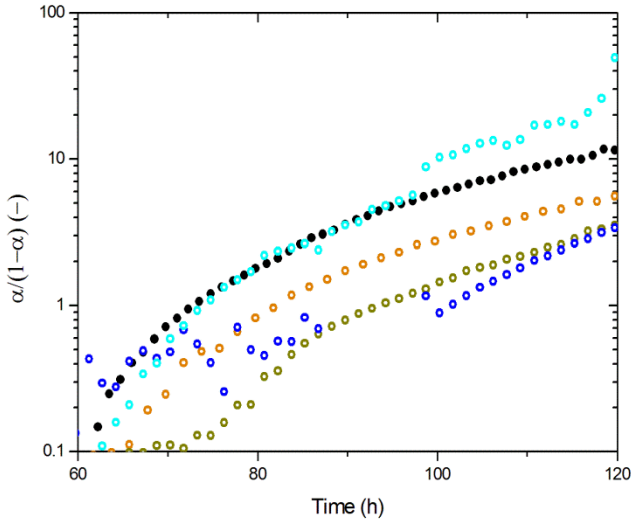
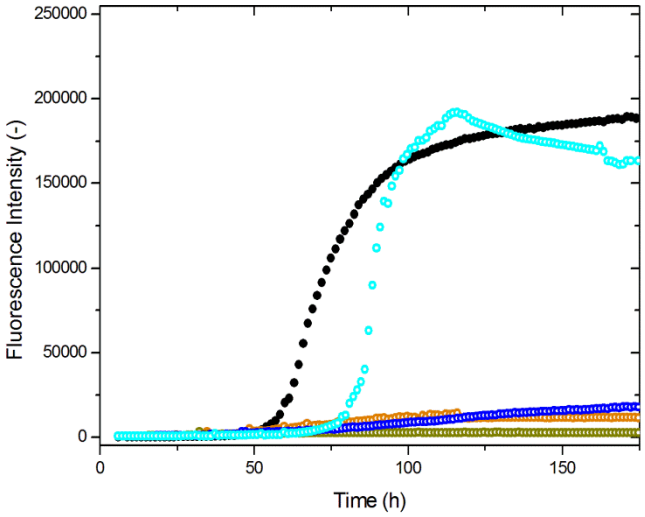


Figure App.1. Influence of benzothiazole on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, $\alpha\,(1-\alpha)$ as a function of time in log-linear scale.

2. Benzothiazole-2-carboxaldehyde



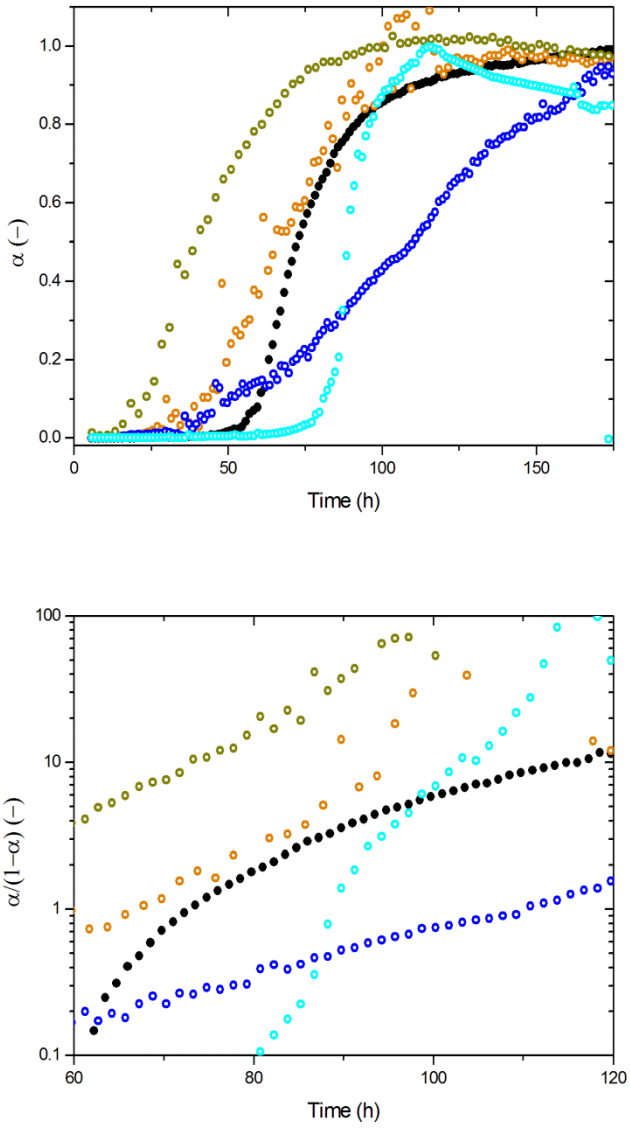
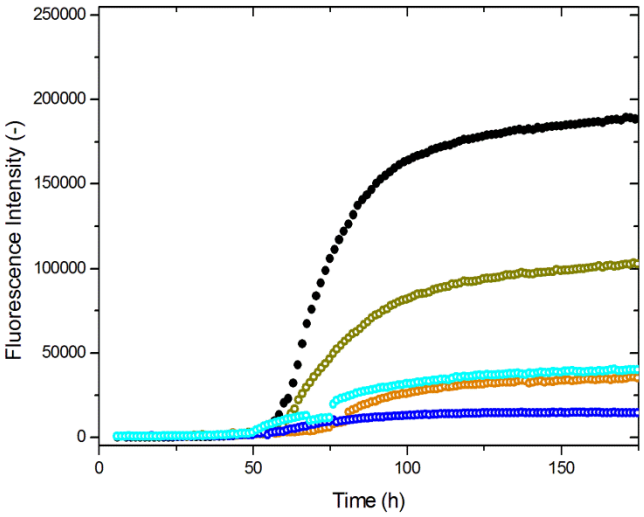
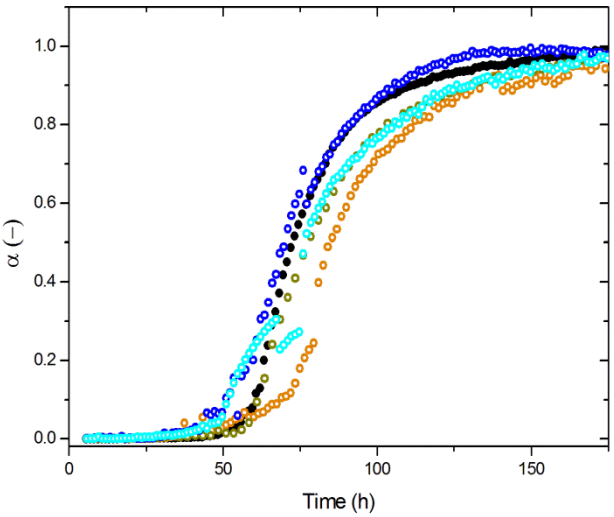


Figure App.2. Influence of benzothiazole-2-carboxaldehyde on the kinetics of aSyn amyloid fibril formation (aSyn 140 μM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 μM (light blue), 50 μM (blue), 100 μM (green) and 200 μM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, $\alpha\,(1-\alpha)$ as a function of time in log-linear scale.

3. Benzothiazole-5-carboxylic acid


Fluorescence Intensity (-)
250000
200000
150000
100000
50000
0
0
50
100
150
Time (h)


α (–)
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
Time (h)

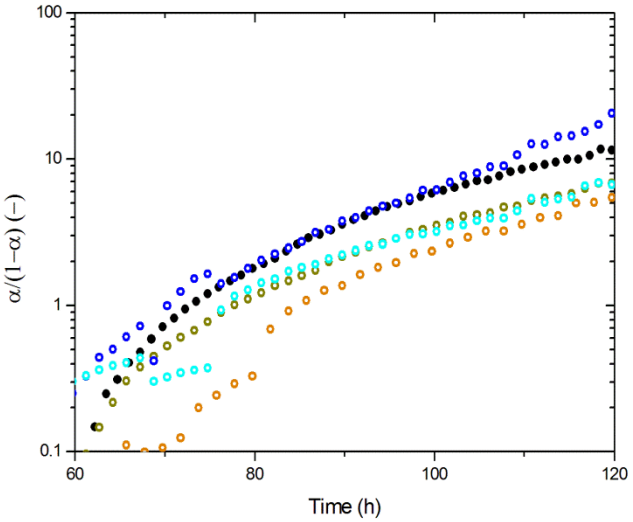
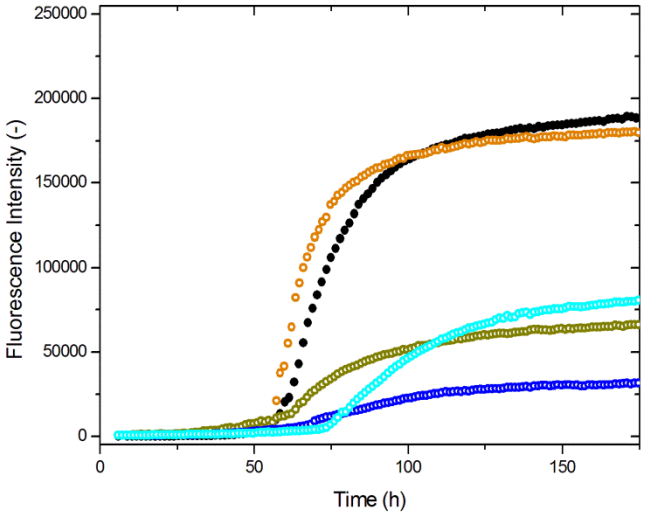
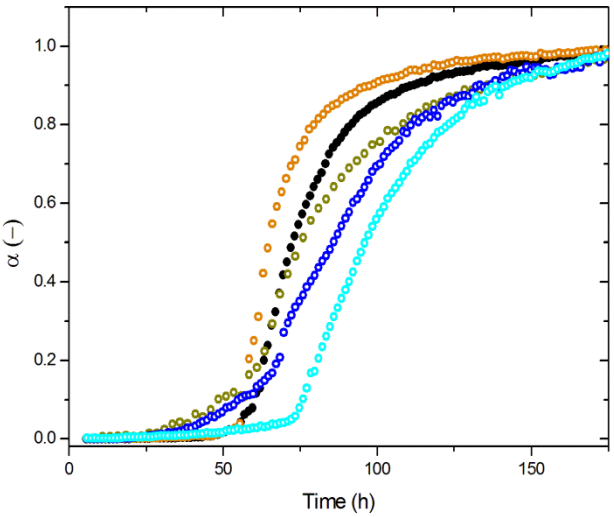

α/(1–α) (–)
100
10
1
0.1
60
80
100
120
Time (h)

Figure App.3. Influence of benzothiazole-5-carboxylic acid on the kinetics of aSyn amyloid fibril formation (aSyn 140 μM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 μM (light blue), 50 μM (blue), 100 μM (green) and 200 μM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, $\alpha\,(1-\alpha)$ as a function of time in log-linear scale.

4. Ethyl benzothiazole-2-carboxylate





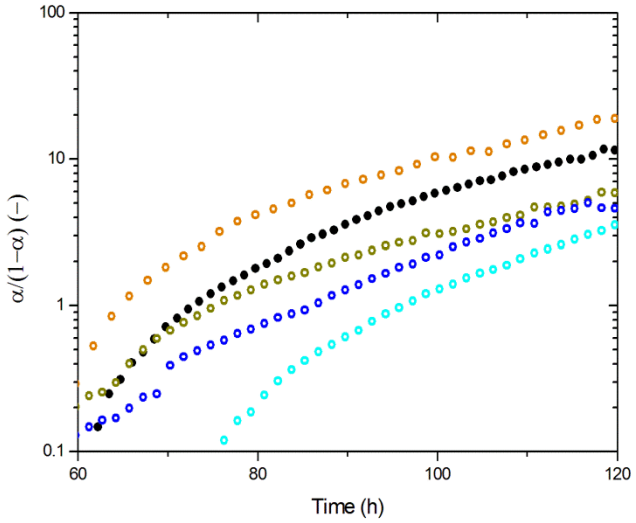
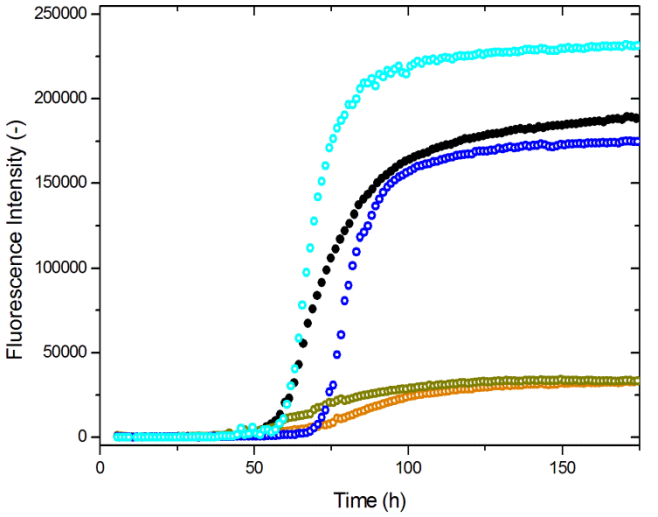


Figure App.4. Influence of ethyl benzothiazole-2-carboxylate on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, $\alpha\,(1-\alpha)$ as a function of time in log-linear scale.

5. 2-(2-Aminophenyl)benzothiazole



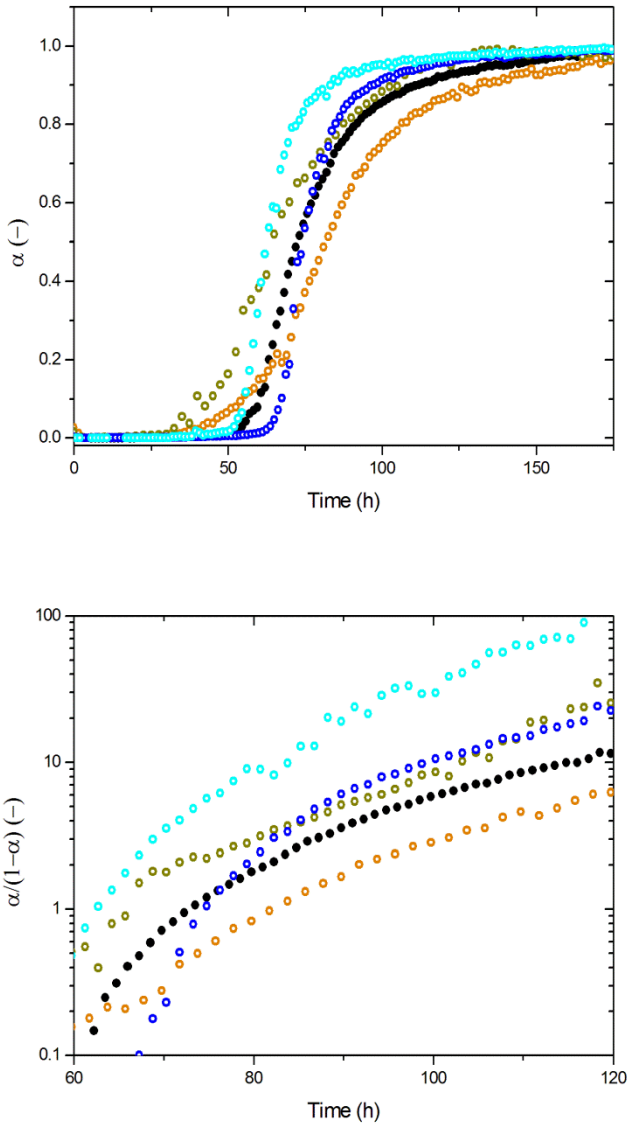
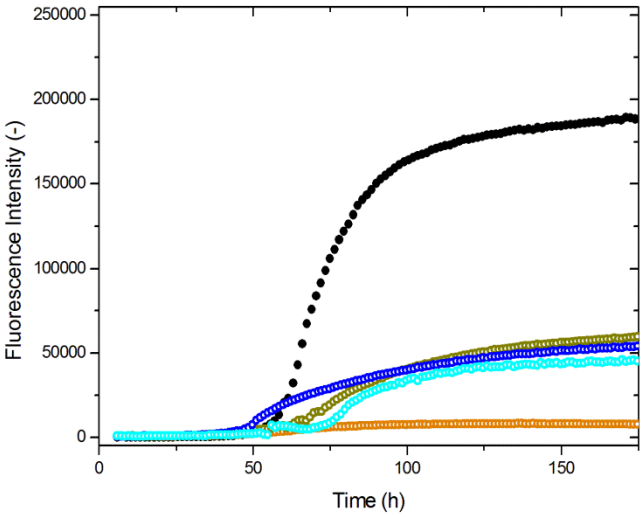
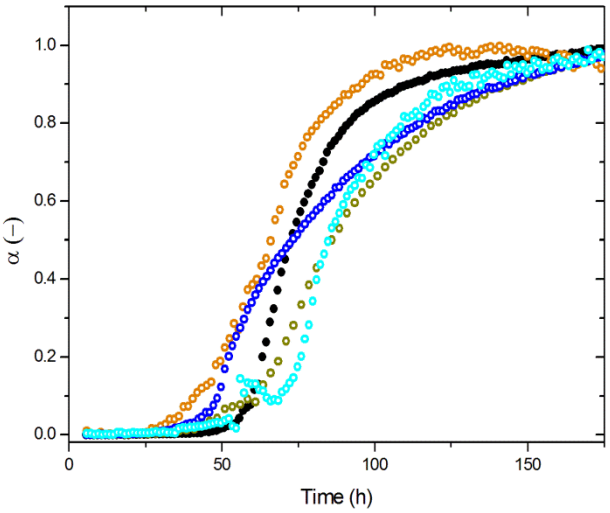


Figure App.5. Influence of 2-(2-aminophenyl)benzothiazole on the kinetics of aSyn amyloid fibril formation (aSyn 140 μM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 μM (light blue), 50 μM (blue), 100 μM (green) and 200 μM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, $\alpha\,(1-\alpha)$ as a function of time in log-linear scale.

6. 2-(4-Aminophenyl)benzothiazole


Fluorescence Intensity (-)
250000
200000
150000
100000
50000
0
0
50
100
150
Time (h)


α (–)
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
Time (h)

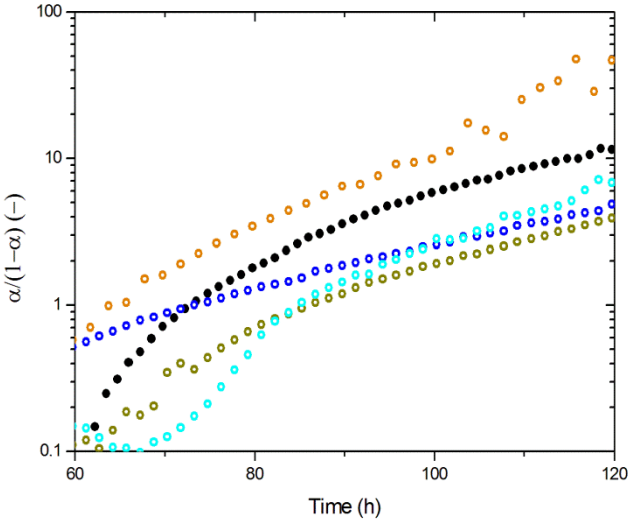
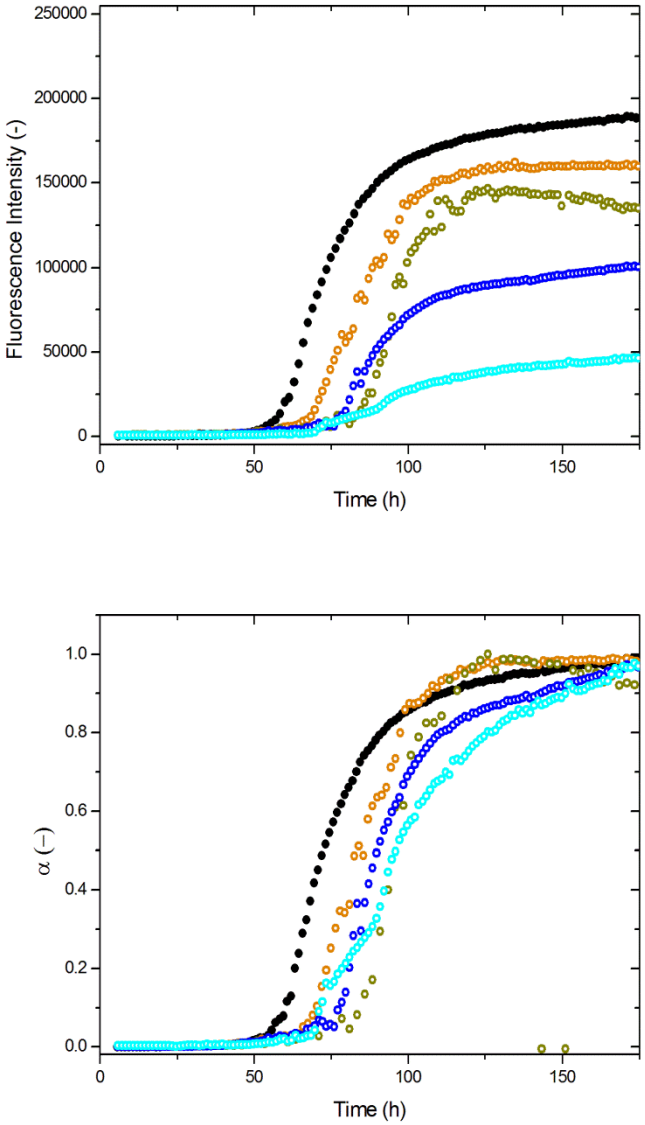

α/(1–α) (–)
100
10
1
0.1
60
80
100
120
Time (h)

Figure App.6. Influence of 2-(4-aminophenyl)benzothiazole on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, $\alpha\,(1-\alpha)$ as a function of time in log-linear scale.

7. 2-(4-Piperidinyl)-1,3-benzothiazole



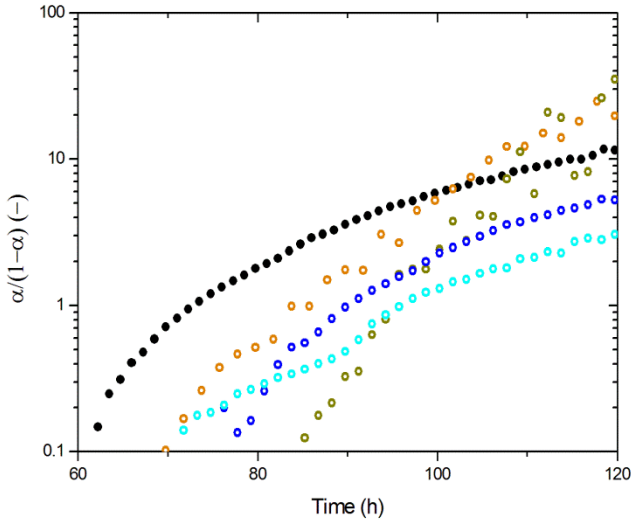
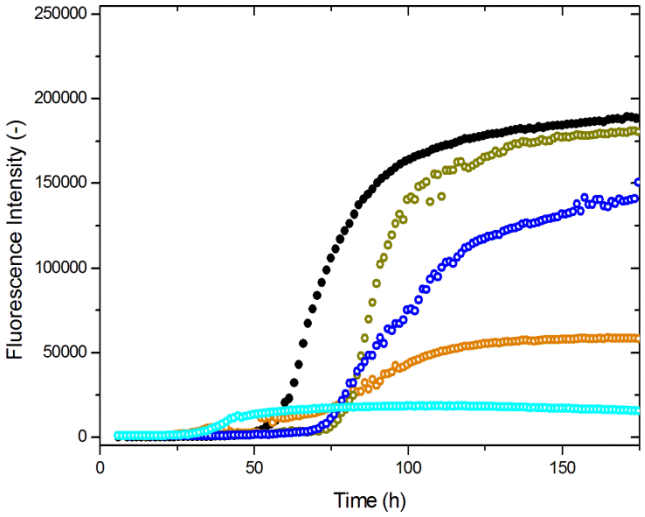


Figure App.7. Influence of 2-(4-piperidinyl)-1,3-benzothiazole on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, α (1−α) as a function of time in log-linear scale.

8. 2-Hydroxy-1,3-benzothiazole-6-carboxylic acid



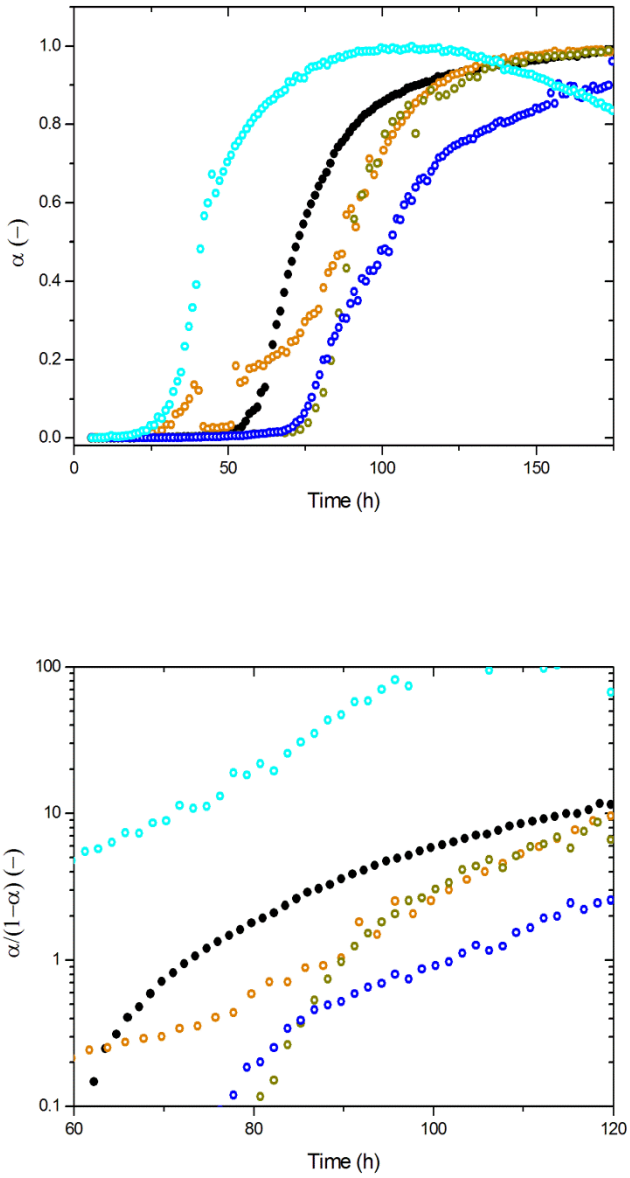
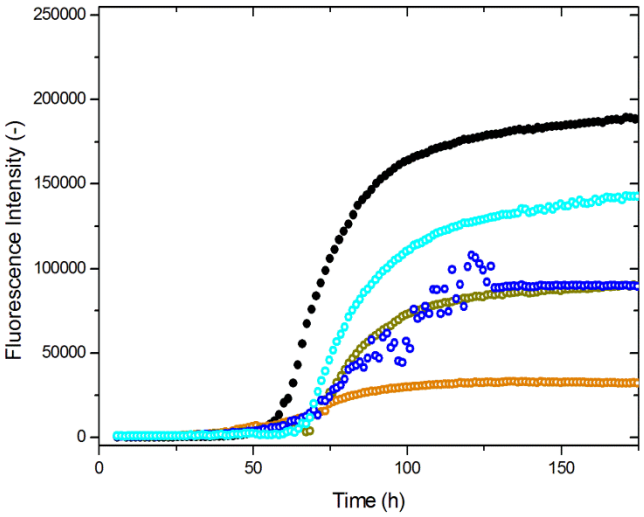
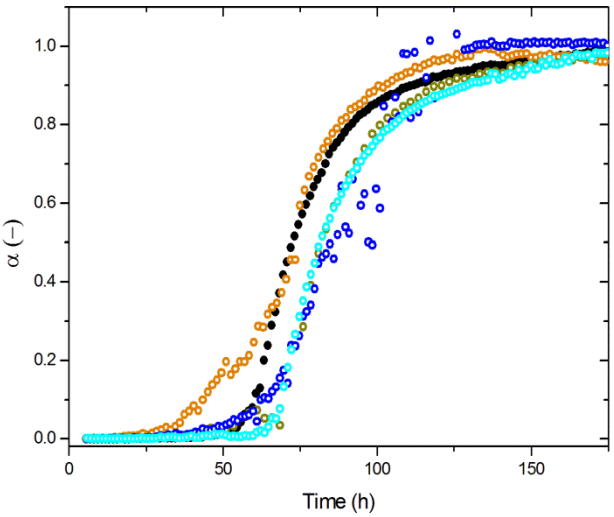


Figure App.8. Influence of 2-hydroxy-1,3-benzothiazole-6-carboxylic acid on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, $\alpha\ (1-\alpha)$ as a function of time in log-linear scale.

9. 2-Methyl-1,3-benzothiazole-6-carboxylic acid


Fluorescence Intensity (-)
250000
200000
150000
100000
50000
0
0
50
100
150
Time (h)


α (–)
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
Time (h)

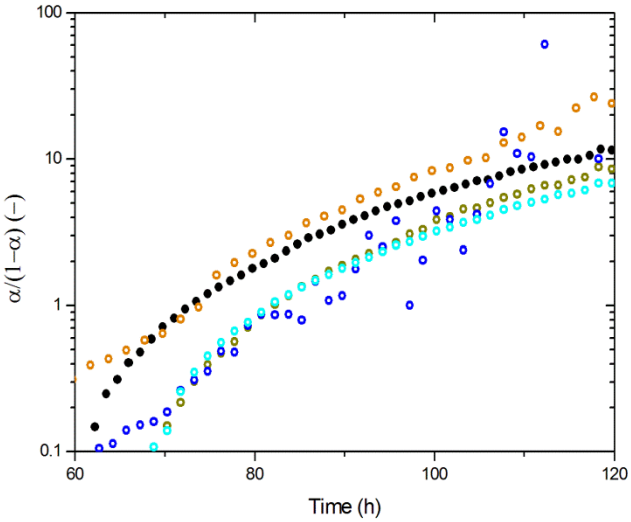
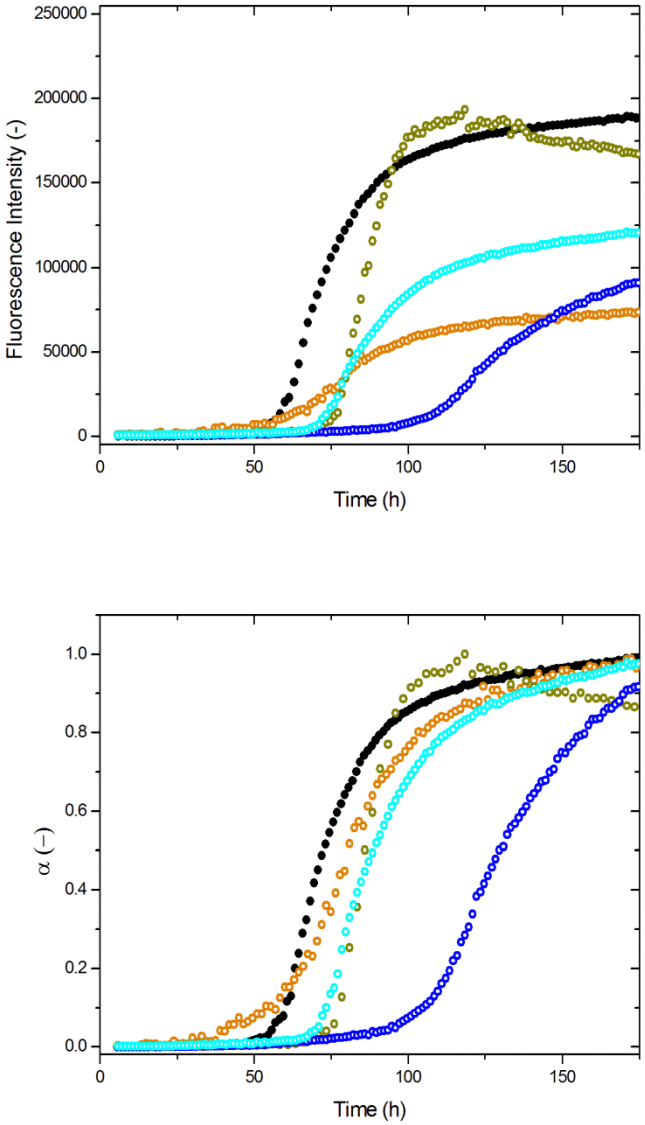

α/(1–α) (–)
100
10
1
0.1
60
80
100
120
Time (h)

Figure App.9. Influence of ethyl 2-methyl-1,3-benzothiazole-6-carboxylic acid on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, $\alpha\ (1-\alpha)$ as a function of time in log-linear scale.

10. 4,5,6,7-Tetrahydro-1,3-benzothiazole-2-carboxylic



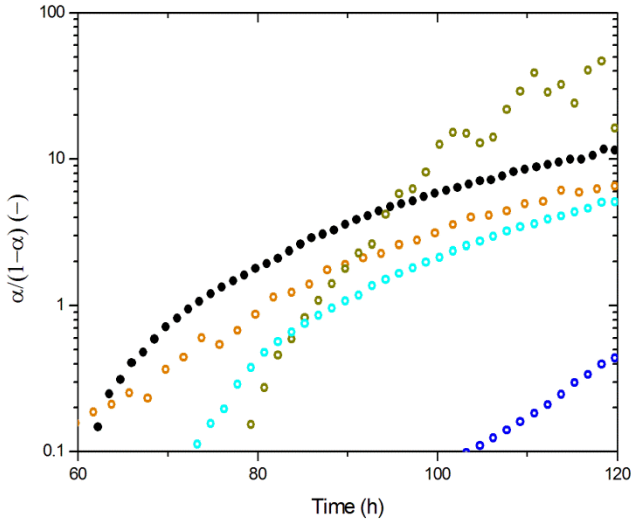
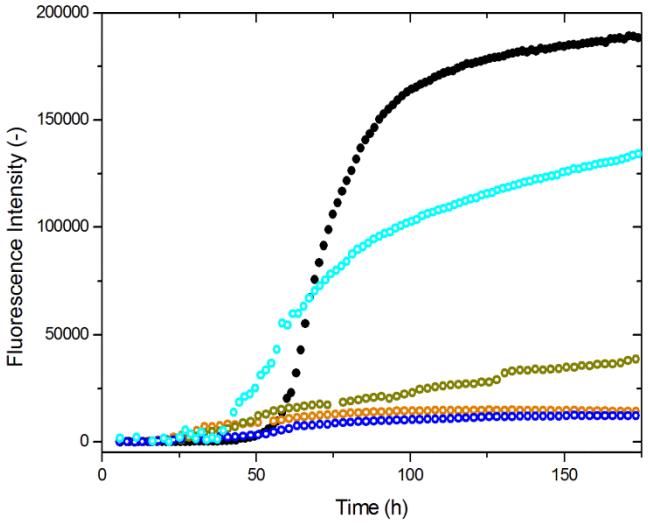


Figure App.10. Influence of ethyl 4,5,6,7-tetrahydro-1,3-benzothiazole-2-carboxylic on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, α (1−α) as a function of time in log-linear scale.

11. Benzothiazole-2-carbonyl chloride



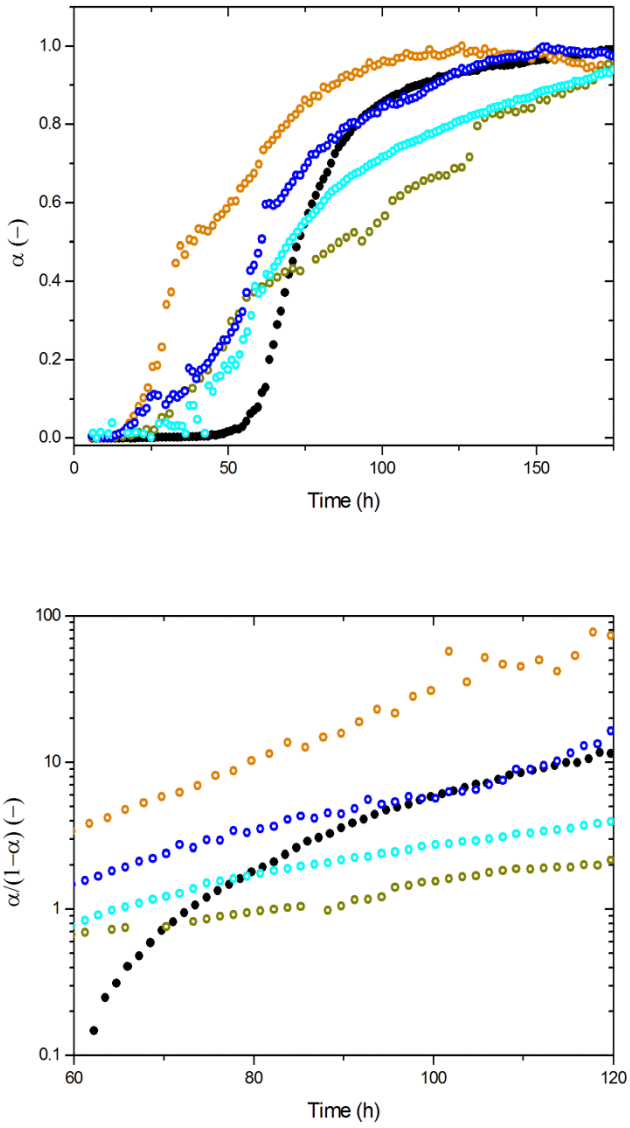
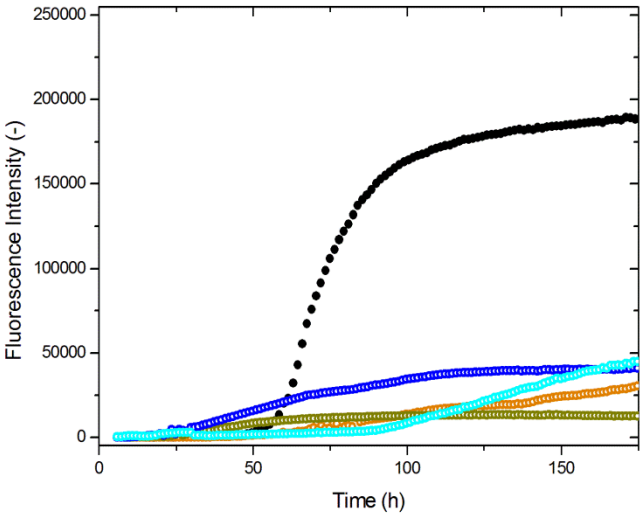
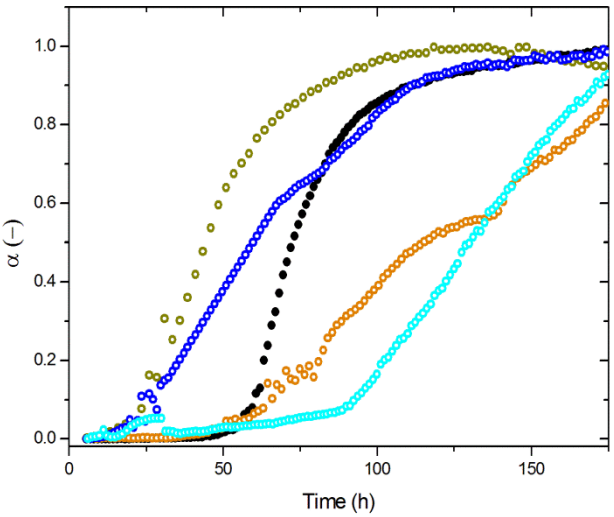


Figure 5.11. Influence of benzothiazole-2-carbonyl chloride on the kinetics of aSyn amyloid fibril formation (aSyn 140 μM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 μM (light blue), 50 μM (blue), 100 μM (green) and 200 μM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, $\alpha\,(1-\alpha)$ as a function of time in log-linear scale.

12. 2-(Chloromethyl)benzothiazole


Fluorescence Intensity (-)
250000
200000
150000
100000
50000
0
0
50
100
150
Time (h)


α (–)
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
Time (h)

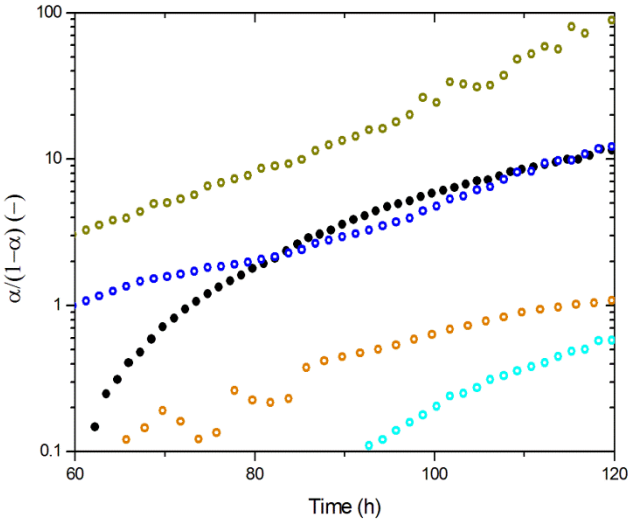
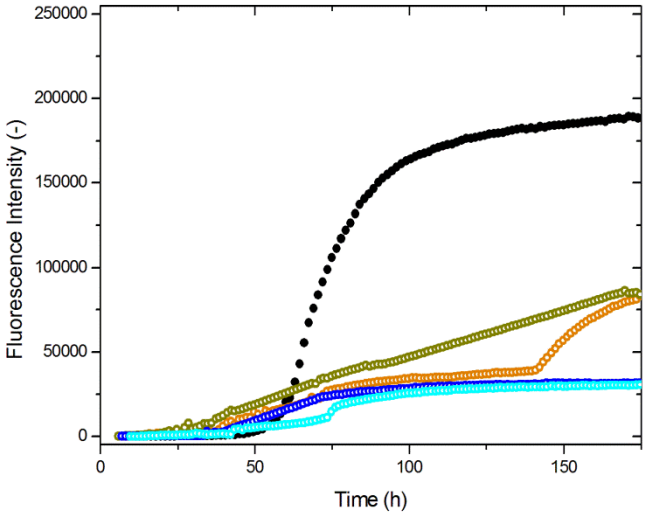
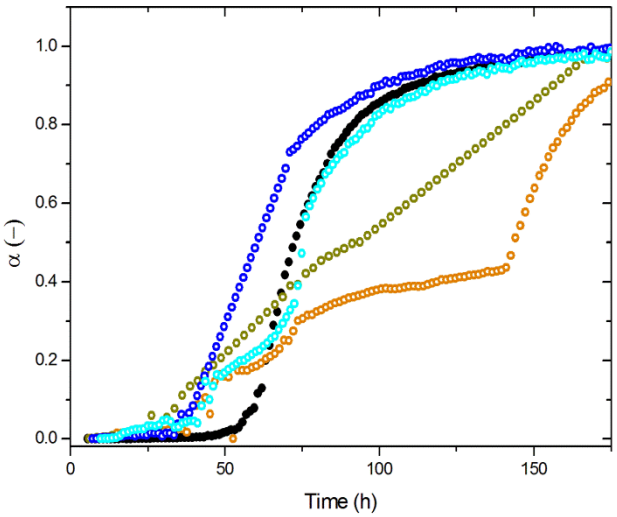

α/(1–α) (–)
100
10
1
0.1
60
80
100
120
Time (h)

Figure App.12. Influence of ethyl 2-(chloromethyl)benzothiazole on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, α $(1-\alpha)$ as a function of time in log-linear scale.

13. 2-(Bromomethyl)-1,3-benzothiazole





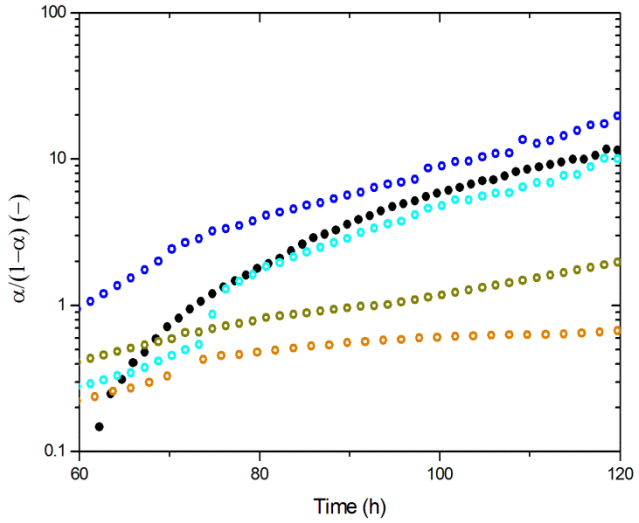
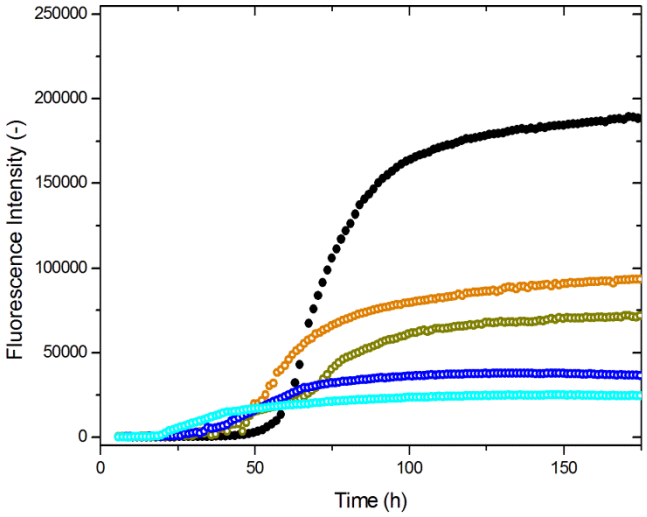


Figure App.13. Influence of ethyl 2-(bromomethyl)-1,3-benzothiazole on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, $\alpha\,(1-\alpha)$ as a function of time in log-linear scale.

14. 2-Chloro-1,3-benzothiazole-6-carboxylic acid



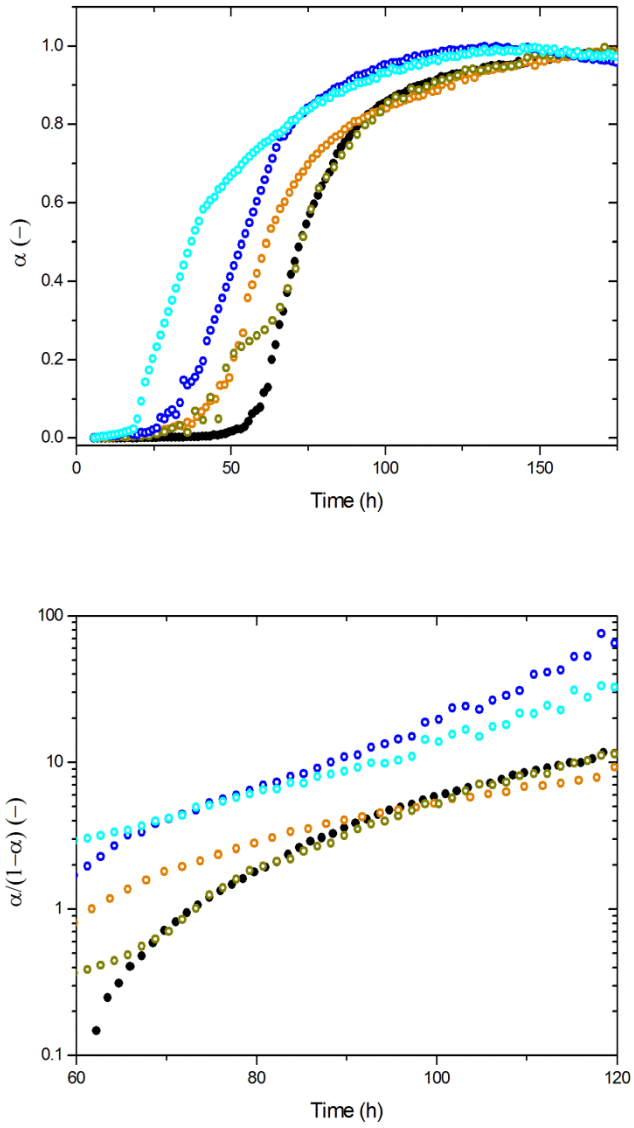
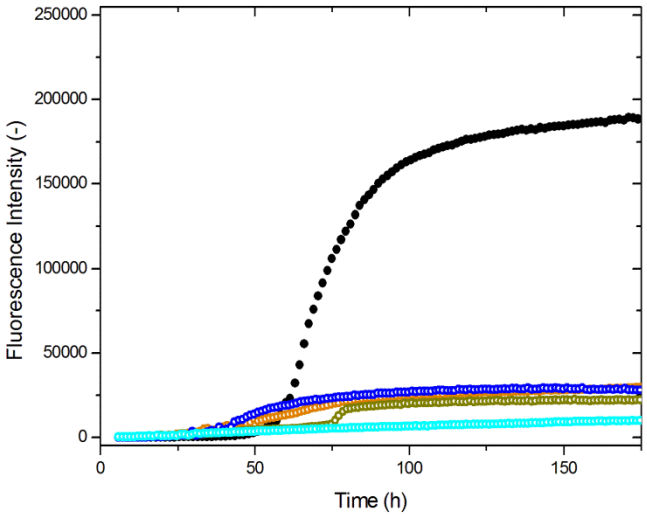
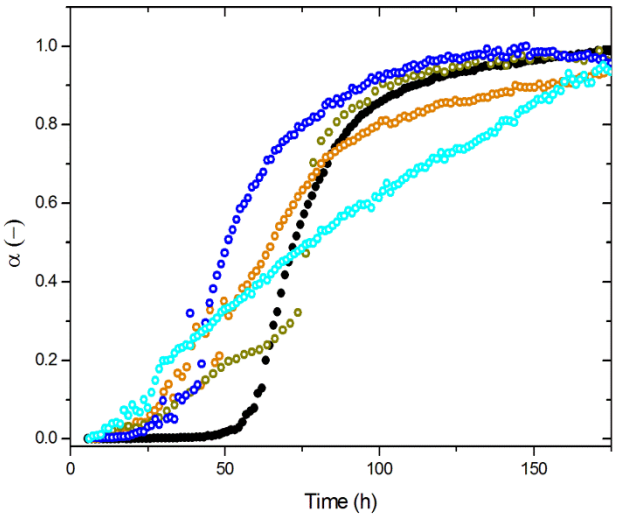


Figure App.14. Influence of ethyl 2-chloro-1,3-benzothiazole-6-carboxylic acid on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, α $(1-\alpha)$ as a function of time in log-linear scale.

15. 6-(Hydroxymethyl)benzothiazole


Fluorescence Intensity (-)
250000
200000
150000
100000
50000
0
0
50
100
150
Time (h)


α (–)
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
Time (h)

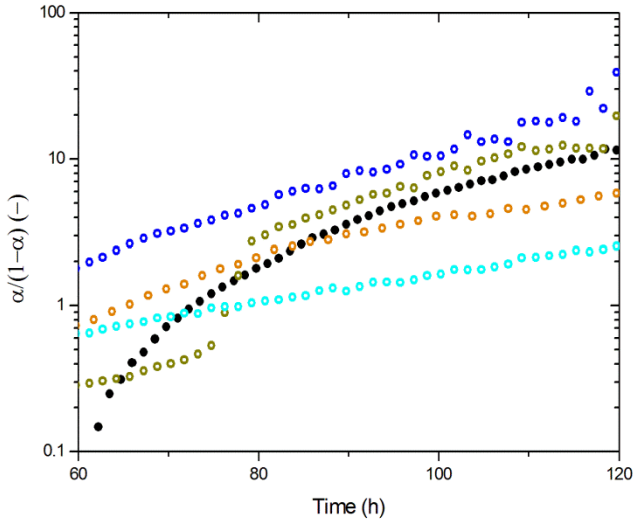
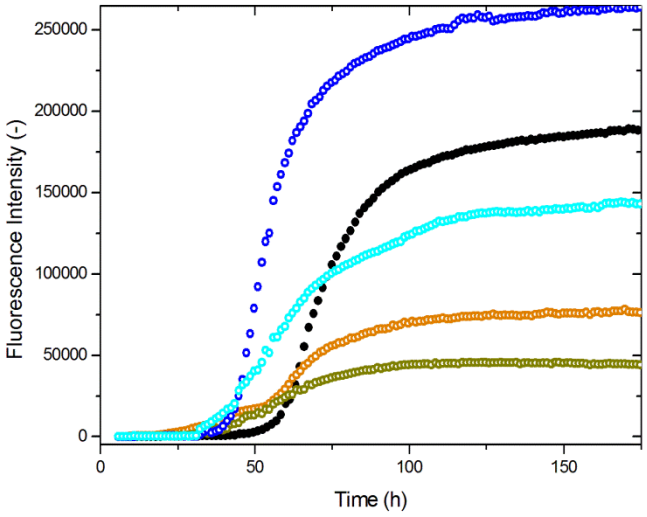
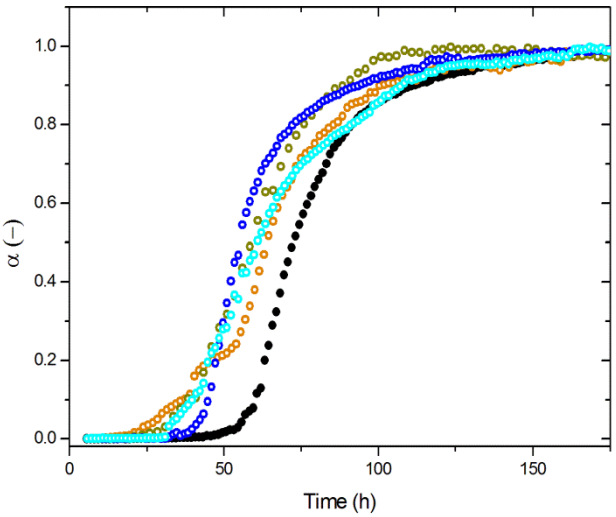

α/(1–α) (–)
100
10
1
0.1
60
80
100
120
Time (h)

Figure App.15. Influence of ethyl 6-(hydroxymethyl)benzothiazole on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, α $(1-\alpha)$ as a function of time in log-linear scale.

16. 2-(1*H*-Pyrazol-3-yl)-1,3-benzothiazole





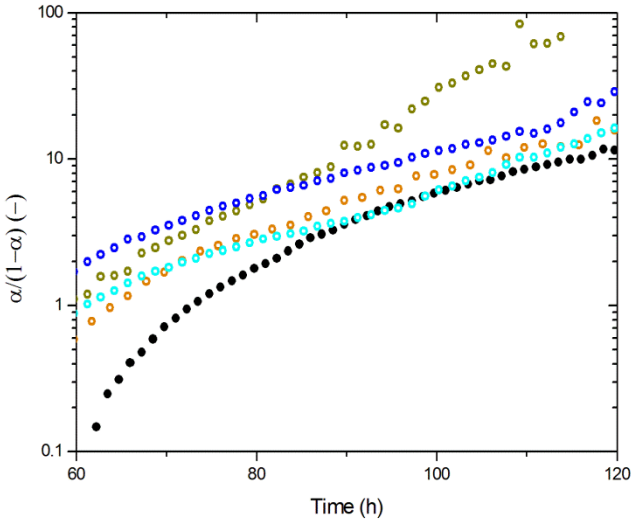
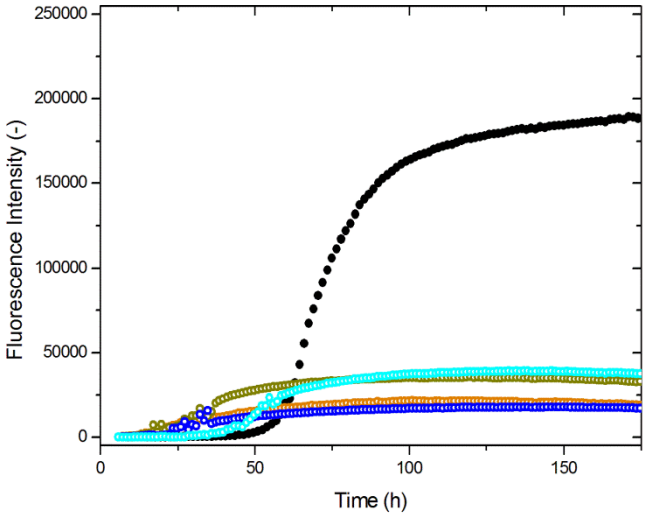


Figure App.16. Influence of 2-(1*H*-pyrazol-3-yl)-1,3-benzothiazole on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, α (1−α) as a function of time in log-linear scale.

17. 2-(1*H*-Pyrrol-1-yl)-1,3-benzothiazole-6-carboxylic



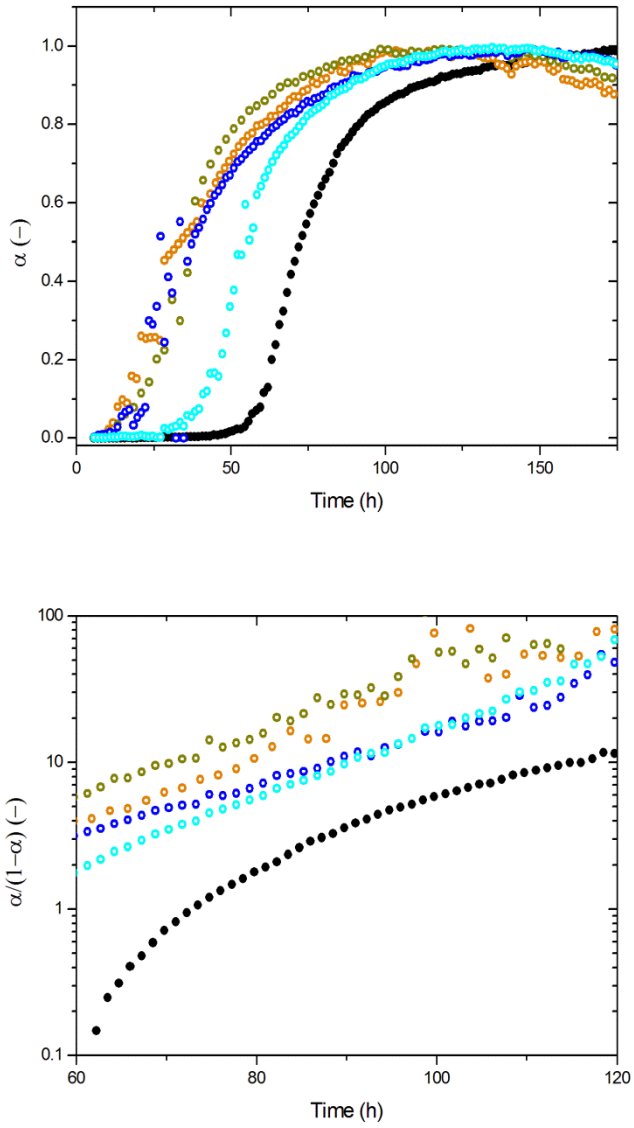
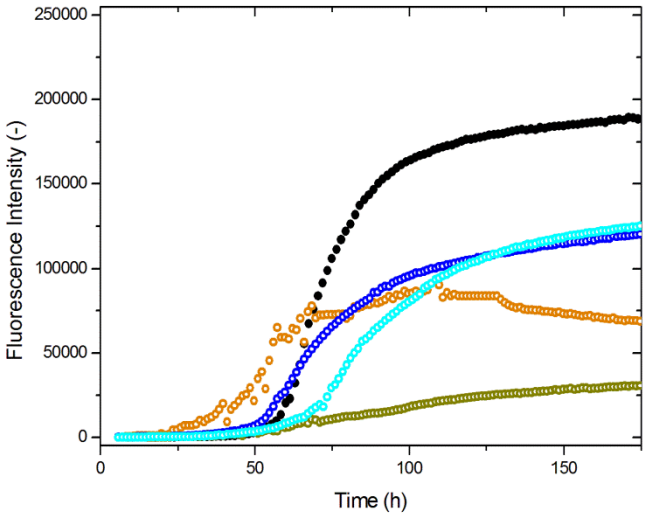
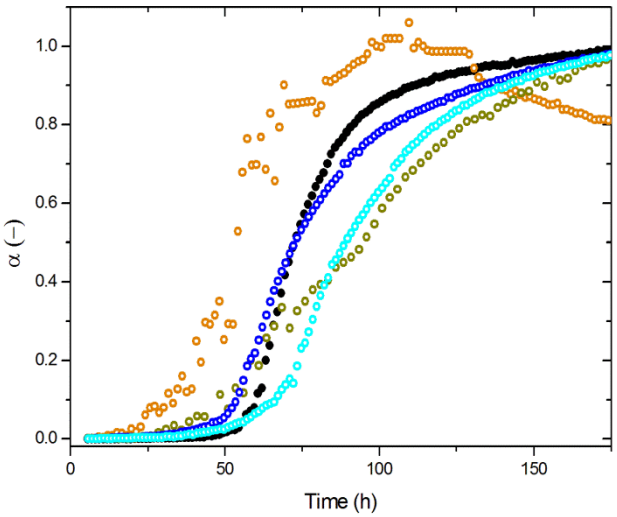


Figure App.17. Influence of 2-(1*H*-pyrrol-1-yl)-1,3-benzothiazole-6-carboxylic on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, α $(1-\alpha)$ as a function of time in log-linear scale.

18. 2-pyrrolidin-2-yl-benzothiazole


Fluorescence Intensity (-)
250000
200000
150000
100000
50000
0
0
50
100
150
Time (h)


α (–)
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
Time (h)

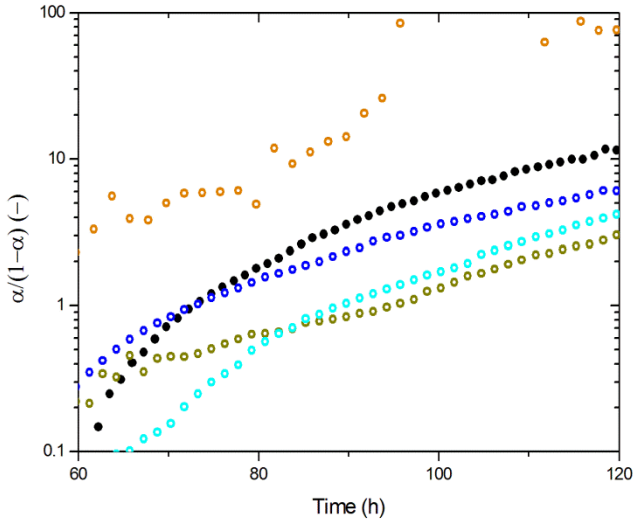

α/(1–α) (–)
100
10
1
0.1
60
80
100
120
Time (h)

Figure App.18. Influence of 2-pyrrolidin-2-yl-benzothiazole on the kinetics of aSyn amyloid fibril formation (aSyn 140 μM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 μM (light blue), 50 μM (blue), 100 μM (green) and 200 μM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, α $(1-\alpha)$ as a function of time in log-linear scale.

19. PGL-135 hydrochloride monohydrate

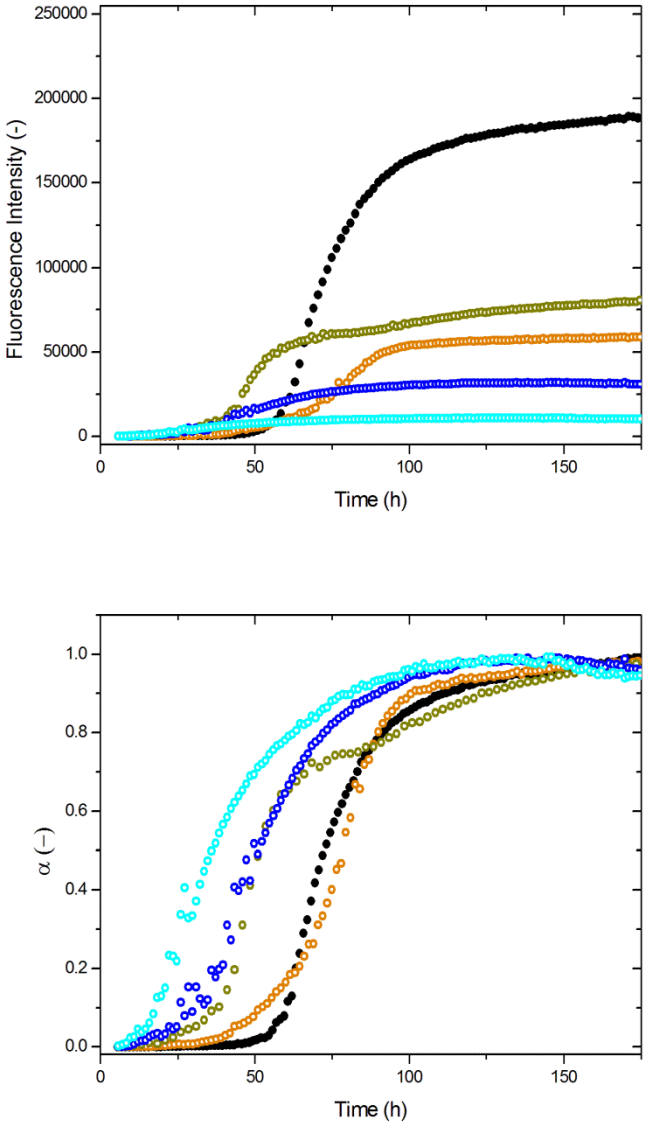
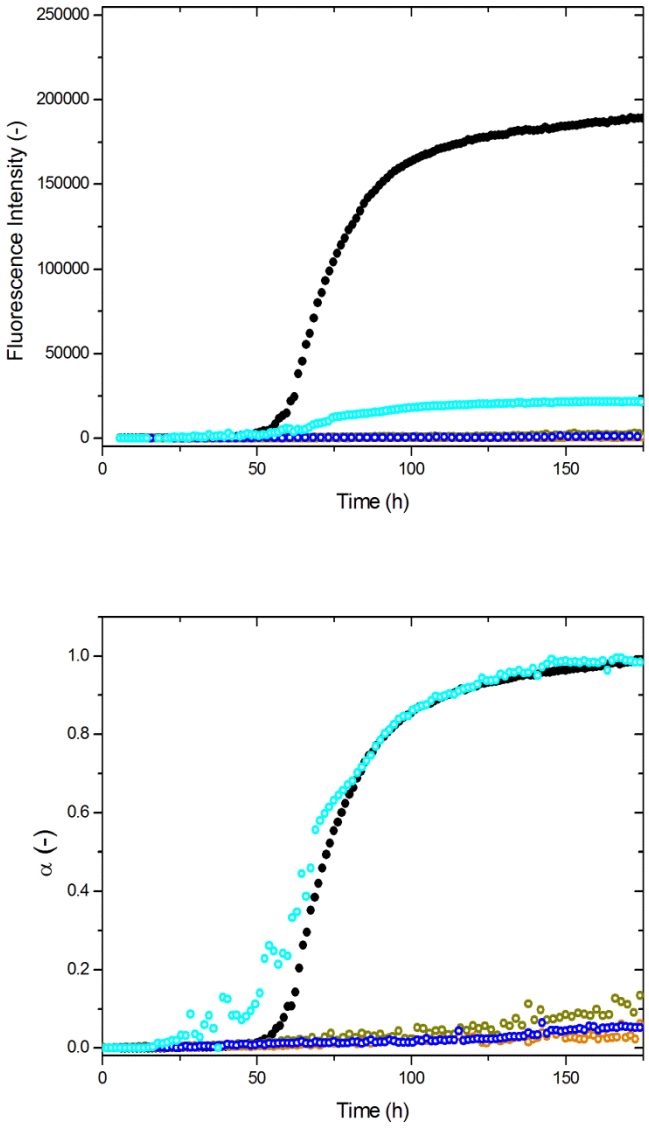


Figure App.19. Influence of PGL-135 hydrochloride monohydrate on the kinetics of aSyn amyloid fibril formation (aSyn 140 μM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 μM (light blue), 50 μM (blue), 100 μM (green) and 200 μM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time.

(Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, $\alpha\,(1-\alpha)$ as a function of time in log-linear scale.

20. Epigallocatechin-3-gallate (EGCG)



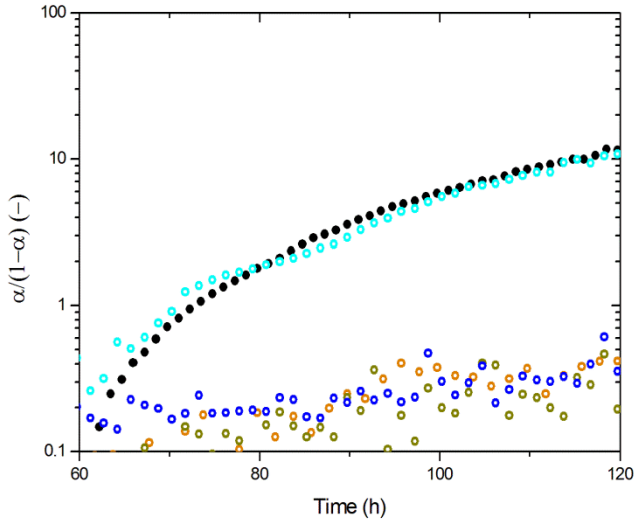
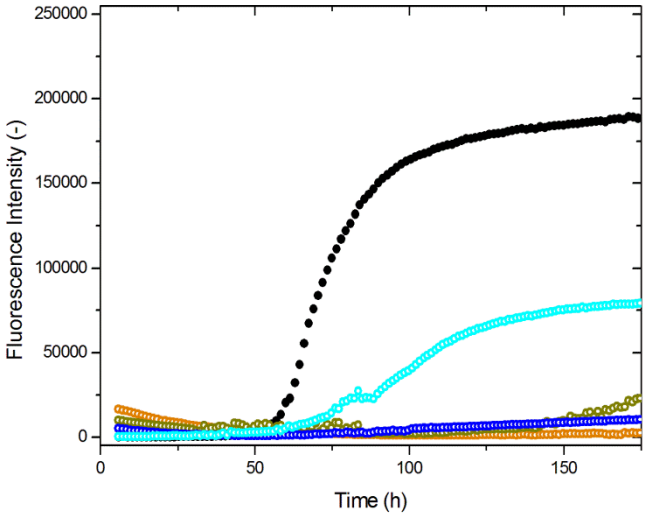


Figure App.20. Influence of EGCG on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, α (1−α) as a function of time in log-linear scale.

21. Curcumin



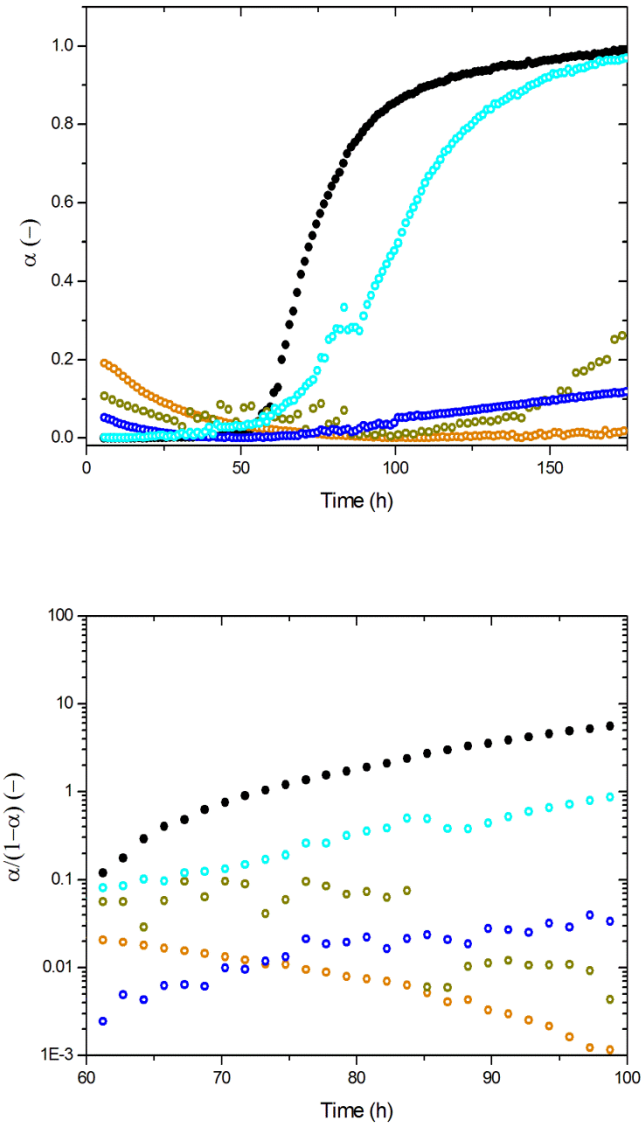


Figure App.21. Influence of curcumin on the kinetics of aSyn amyloid fibril formation (aSyn 140 µM, pH 7.4 and 37 ºC). 4 different compound concentrations were assayed: 5 µM (light blue), 50 µM (blue), 100 µM (green) and 200 µM (orange). aSyn in the absence of compound (black). (Top) ThT fluorescence increase represented as a function of the incubation time. (Middle) ThT fluorescence increase represented in normalized units as a function of the incubation time. (Bottom) Representation in modified coordinates, α (1–α) as a function of time in log-linear scale.

Bibliography

1. Adachi, M., So, M., Sakurai, K., Kardos, J. & Goto, Y. Supersaturation-limited and Unlimited Phase Transitions Compete to Produce the Pathway Complexity in Amyloid Fibrillation. *J. Biol. Chem.* **290,** 18134–45 (2015).

2. Durbin, S. D. & Feher, G. Protein crystallization. *Annu. Rev. Phys. Chem.* **47,** 171–204 (1996).

3. Yoshimura, Y. *et al.* Distinguishing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 14446–51 (2012).

4. Kitayama, H. *et al.* A common mechanism underlying amyloid fibrillation and protein crystallization revealed by the effects of ultrasonication. *Biochim. Biophys. Acta* **1834,** 2640–6 (2013).

5. Crespo, R., Martins, P. M., Gales, L., Rocha, F. & Damas, A. M. Potential use of ultrasound to promote protein crystallization. *J. Appl. Crystallogr.* **43,** 1419–1425 (2010).

6. Crespo, R., Rocha, F. a, Damas, A. M. & Martins, P. M. A generic crystallization-like model that describes the kinetics of amyloid fibril formation. *J. Biol. Chem.* **287,** 30585–94 (2012).

7. Ferreira, C. *et al.* Small temperature oscillations promote protein crystallization. *CrystEngComm* **13,** 3051 (2011).

8. Ikenoue, T. *et al.* Heat of supersaturation-limited amyloid burst directly monitored by isothermal titration calorimetry. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 6654–9 (2014).

9. Mulaj, M., Foley, J. & Muschol, M. Amyloid Oligomers and Proto fi brils, but Not Filaments, Self- Replicate from Native Lysozyme. (2015).

10. Muta, H. *et al.* Supersaturation-limited Amyloid fibrillation of Insulin revealed by ultrasonication. *J. Biol. Chem.* **289,** 18228–18238 (2014).

11. Yagi, H., Abe, Y., Takayanagi, N. & Goto, Y. Elongation of amyloid fibrils through lateral binding of monomers revealed by total internal reflection fluorescence microscopy. *Biochim. Biophys. Acta* **1844,** 1881–8 (2014).

12. Ciryam, P., Kundra, R., Morimoto, R. I., Dobson, C. M. & Vendruscolo, M. Supersaturation is a major driving force for protein aggregation in neurodegenerative diseases. *Trends Pharmacol. Sci.* **36,** 72–7 (2015).

13. Hall, D., Kardos, J., Edskes, H., Carver, J. a & Goto, Y. A multi-pathway perspective on protein aggregation: implications for control of the rate and extent of amyloid formation. *FEBS Lett.* **589,** 672–679 (2015).

14. Westermark, P. E. R. *et al.* Amyloid : Toward terminology clarification Report from the Nomenclature Committee of the International Society of Amyloidosis. *Amyloid* **12,** 1–4 (2005).

15. Sipe, J. D. & Cohen, A. S. Review : History of the Amyloid Fibril. *J. Struct. Biol.* **130,** 88–98 (2000).

16. Rudolf Virchow. Zur Cellulose-Frage. *Arch Pathol Anat* **6,** 416–426 (1984).

17. Rambaran, R. N. & Serpell, L. C. Amyloid fibrils: abnormal protein assembly. *Prion* **2,** 112–117 (2008).

18. Pepys, M. B. Pathogenesis, diagnosis and treatment of systemic amyloidosis. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **356,** 203–10; discussion 210–1 (2001).

19. Buxbaum, J. N. Diseases of protein conformation: What do in vitro experiments tell us about in vivo diseases? *Trends Biochem. Sci.* **28,** 585–592 (2003).

20. Chiti, F. *et al.* Designing conditions for in vitro formation of amyloid protofilaments and fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 3590–3594 (1999).

21. Schmittschmitt, J. P. & Scholtz, J. M. The role of protein stability, solubility, and net charge in amyloid fibril formation. *Protein Sci.* **12,** 2374–2378 (2003).

22. Dobson, C. M. Protein misfolding, evolution and disease. *Trends Biochem. Sci.* **24,** 329–332 (1999).

23. Fändrich, M. On the structural definition of amyloid fibrils and other polypeptide aggregates. *Cell. Mol. Life Sci.* **64,** 2066–78 (2007).

24. Chen, B., Thurber, K. R., Shewmaker, F., Wickner, R. B. & Tycko, R. Measurement of amyloid fibril mass-per-length by tilted-beam transmission electron microscopy. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 14339–44 (2009).

25. Nilsson, M. R. Techniques to study amyloid fibril formation in vitro. *Methods* **34,** 151–60 (2004).

26. Knowles, T. P. J. *et al.* Kinetics and thermodynamics of amyloid formation from direct measurements of fluctuations in fibril mass. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 10016–21 (2007).

27. Lednev, I. K. Amyloid fibrils: The eighth wonder of the world in protein folding and aggregation. *Biophys. J.* **106,** 1433–1435 (2014).

28. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921 (2001).

29. Eisele, Y. S. *et al.* Targeting protein aggregation for the treatment of degenerative diseases. *Nat. Rev. Drug Discov.* (2015). doi:10.1038/nrd4593

30. Kelly, J. W. Towards an understanding of amyloidogenesis. *Nat. Struct. Biol.* **9,** 323–325 (2002).

31. Stefani, M. & Rigacci, S. Protein folding and aggregation into amyloid: the interference by natural phenolic compounds. *Int. J. Mol. Sci.* **14,** 12411–57 (2013).

32. Kamihira, M., Naito, a, Tuzi, S., Nosaka, a Y. & Saitô, H. Conformational transitions and fibrillation mechanism of human calcitonin as studied by high-resolution solid-state 13C NMR. *Protein Sci.* **9,** 867–77 (2000).

33. Moniatte, M., van der Goot, F. G., Buckley, J. T., Pattus, F. & van Dorsselaer, a. Characterisation of the heptameric pore-forming complex of the Aeromonas toxin aerolysin using MALDI-TOF mass spectrometry. *FEBS Lett.* **384,** 269–272 (1996).

34. Gillam, J. E. & MacPhee, C. E. Modelling amyloid fibril formation kinetics: mechanisms of nucleation and growth. *J. Phys. Condens. Matter* **25,** 373101 (2013).

35. Kodaka, M. Requirements for generating sigmoidal time-course aggregation in nucleation-dependent polymerization model. *Biophys. Chem.* **107,** 243–53 (2004).

36. Jarrett, J. T. & Lansbury, P. T. Amyloid fibril formation requires a chemically discriminating nucleation event: studies of an amyloidogenic sequence from the bacterial protein OsmB. *Biochemistry* **31,** 12345–52 (1992).

37. Kodaka, M. Interpretation of concentration-dependence in aggregation kinetics. *Biophys. Chem.* **109,** 325–32 (2004).

38. Li, H., Rahimi, F., Sinha, S., Maiti, P. & Bitan, G. Amyloids and Protein Aggregation – Analytical Methods. *Encyclopedia of Analytical Chemistry* 1–32 (2009).

39. Goldsbury, C., Kistler, J., Aebi, U., Arvinte, T. & Cooper, G. J. Watching amyloid fibrils grow by time-lapse atomic force microscopy. *J. Mol. Biol.* **285,** 33–9 (1999).

40. Jarrett, J. T. & Lansbury Jr., P. T. Seeding 'one-dimensional crystallization' of amyloid: A pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* **73,** 1055–1058 (1993).

41. Fezoui, Y. & Teplow, D. B. Kinetic studies of amyloid beta-protein fibril assembly. Differential effects of alpha-helix stabilization. *J. Biol. Chem.* **277,** 36948–54 (2002).

42. Harper, J. D. & Lansbury, P. T. Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annu. Rev. Biochem.* **66,** 385–407 (1997).

43. Frieden, C. Protein aggregation processes : In search of the mechanism. 2334–2344 (2007). doi:10.1110/ps.073164107.tems

44. Biancalana, M. & Koide, S. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochim. Biophys. Acta* **1804,** 1405–12 (2010).

45. Levine, H. Thioflavine T interaction with synthetic Alzheimer ' s disease @ -amyloid peptides : Detection of amyloid aggregation in solution. 404–410 (1993).

46. Lindgren, M., Sörgjerd, K. & Hammarström, P. Detection and characterization of aggregates, prefibrillar amyloidogenic oligomers, and protofibrils using

fluorescence spectroscopy. *Biophys. J.* **88,** 4200–4212 (2005).

47. Howie, A. J., Brewer, D. B., Howell, D. & Jones, A. P. Physical basis of colors seen in Congo red-stained amyloid in polarized light. *Lab. Invest.* **88,** 232–242 (2008).

48. Groenning, M. *et al.* Binding mode of Thioflavin T in insulin amyloid fibrils. *J. Struct. Biol.* **159,** 483–497 (2007).

49. Klunk, W. E. *et al.* Imaging Brain Amyloid in Alzheimer's Disease with Pittsburgh Compound-B. *Ann. Neurol.* **55,** 306–319 (2004).

50. Romano, M. & Buratti, E. Florbetapir F 18 for brain imaging of β-amyloid plaques. *Drugs Today (Barc).* **49,** 181–93 (2013).

51. Morris, A. M., Watzky, M. a & Finke, R. G. Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature. *Biochim. Biophys. Acta* **1794,** 375–97 (2009).

52. Ferrone, F. a, Hofrichter, J., Sunshine, H. R. & Eaton, W. a. Kinetic studies on photolysis-induced gelation of sickle cell hemoglobin suggest a new mechanism. *Biophys. J.* **32,** 361–380 (1980).

53. Ruschak, A. M. & Miranker, A. D. Fiber-dependent amyloid formation as catalysis of an existing reaction pathway. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 12341–6 (2007).

54. Hong, L. & Yong, W.-A. Simple moment-closure model for the self-assembly of breakable amyloid filaments. *Biophys. J.* **104,** 533–40 (2013).

55. Ferrone, F. Analysis of protein aggregation kinetics. *Methods Enzymol.* **309,** 256–274 (1999).

56. Knowles, T. P. J. *et al.* An analytical solution to the kinetics of breakable filament assembly. *Science* **326,** 1533–1537 (2009).

57. Watzky, M. A. & Finke, R. G. Transition metal nanocluster formation kinetic and mechanistic studies. A new mechanism when hydrogen is the reductant: Slow, continuous nucleation and fast autocatalytic surface growth. *J. Am. Chem. Soc.* **119,** 10382–10400 (1997).

58. Rochet, J.-C. Novel therapeutic strategies for the treatment of protein-misfolding diseases. *Expert Rev. Mol. Med.* **9,** 1–34 (2007).

59. Lee, V. M. Y. Amyloid binding ligands as Alzheimer's disease therapies. *Neurobiol. Aging* **23,** 1039–1042 (2002).

60. Porat, Y., Abramowitz, A. & Gazit, E. Inhibition of amyloid fibril formation by polyphenols: Structural similarity and aromatic interactions as a common inhibition mechanism. *Chem. Biol. Drug Des.* **67,** 27–37 (2006).

61. Sunde, M. *et al.* Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* **273,** 729–39 (1997).

62. Kelly, J. W. Amyloid fibril formation and protein misassembly: a structural quest for insights into amyloid and prion diseases. *Structure* **5,** 595–600 (1997).

63. Greenwald, J. & Riek, R. Biology of amyloid: structure, function, and regulation. *Structure* **18,** 1244–60 (2010).

64. Roychaudhuri, R., Yang, M., Hoshi, M. M. & Teplow, D. B. Amyloid beta-protein assembly and Alzheimer disease. *J. Biol. Chem.* **284,** 4749–4753 (2009).

65. Bhak, G., Choe, Y. J. & Paik, S. R. Mechanism of amyloidogenesis: Nucleation-dependent fibrillation versus double-concerted fibrillation. *BMB Rep.* **42,** 541–551 (2009).

66. Bondos, S. E. Methods for measuring protein aggregation. *Curr. Anal. Chem.* **2,** 157–170 (2006).

67. Naiki, H., Higuchi, K., Hosokawa, M. & Takeda, T. Fluorometric determination of amyloid fibrils in vitro using the fluorescent dye, thioflavine T. *Anal. Biochem.* **177,** 244–249 (1989).

68. Xue, W.-F., Homans, S. W. & Radford, S. E. Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 8926–31 (2008).

69. Naiki, H., Hasegawa, K., Yamaguchi, I., Nakamura, H. & Gejyo, F. Apolipoprotein E and Antioxidants Have Different Mechanisms of Inhibiting Alzheimer ' s - Amyloid Fibril Formation in Vitro. *Biochemistry* **37,** 17882–17889 (1998).

70. Hurshman, A. R., White, J. T., Powers, E. T. & Kelly, J. W. Transthyretin aggregation under partially denaturing conditions is a downhill polymerization. *Biochemistry* **43,** 7365–81 (2004).

71. Naiki, H., Gejyo, F. & Nakakuki, K. Concentration-dependent inhibitory effects of apolipoprotein E on Alzheimer's β-amyloid fibril formation in vitro. *Biochemistry* **36,** 6243–6250 (1997).

72. Watzky, M. A., Morris, A. M., Ross, E. D. & Finke, R. G. Fitting yeast and mammalian prion aggregation kinetic data with the finke-watzky two-step model of nucleation and autocatalytic growth. *Biochemistry* **47,** 10790–10800 (2008).

73. Schmidt, H., Madsen, M. F., Danø, S. & Cedersund, G. Complexity reduction of biochemical rate expressions. *Bioinformatics* **24,** 848–54 (2008).

74. Bernacki, J. P. & Murphy, R. M. Model discrimination and mechanistic interpretation of kinetic data in protein aggregation studies. *Biophys. J.* **96,** 2871–87 (2009).

75. Morris, A. M., Watzky, M. a, Agar, J. N. & Finke, R. G. Fitting neurological protein aggregation kinetic data via a 2-step, minimal/'Ockham's razor' model: the Finke-Watzky mechanism of nucleation followed by autocatalytic surface growth. *Biochemistry* **47,** 2413–27 (2008).

76. Fernández, C. O. *et al.* NMR of alpha-synuclein-polyamine complexes elucidates the mechanism and kinetics of induced aggregation. *EMBO J.* **23,** 2039–46 (2004).

77. Sluzky, V., Tamada, J. a, Klibanov, a M. & Langer, R. Kinetics of insulin aggregation in aqueous solutions upon agitation in the presence of hydrophobic surfaces. *Proc. Natl. Acad. Sci. U. S. A.* **88,** 9377–81 (1991).

78. O'Nuallain, B., Shivaprasad, S., Kheterpal, I. & Wetzel, R. Thermodynamics of A beta(1-40) amyloid fibril elongation. *Biochemistry* **44,** 12709–18 (2005).

79. Hasegawa, K., Ono, K., Yamada, M. & Naiki, H. Kinetic modeling and determination of reaction constants of Alzheimer's β-amyloid fibril extension and dissociation

using surface plasmon resonance. *Biochemistry* **41,** 13489–13498 (2002).

80. Auer, S. & Kashchiev, D. Insight into the correlation between lag time and aggregation rate in the kinetics of protein aggregation. *Proteins* **78,** 2412–6 (2010).

81. Jarrett, J. T., Berger, E. P. & Lansbury, P. T. The carboxy terminus of the .beta. amyloid protein is critical for the seeding of amyloid formation: Implications for the pathogenesis of Alzheimer's disease. *Biochemistry* **32,** 4693–4697 (1993).

82. Markov, I. V. *Crystal Growth for Begginers*. (World Scientific Publishing, 2003).

83. Minton, a P. Implications of macromolecular crowding for protein assembly. *Curr. Opin. Struct. Biol.* **10,** 34–9 (2000).

84. Chiti, F. & Dobson, C. M. Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol.* **5,** 15–22 (2009).

85. Stöhr, J. *et al.* Mechanisms of prion protein assembly into amyloid. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 2409–14 (2008).

86. Bhak, G., Lee, J.-H., Hahn, J.-S. & Paik, S. R. Granular assembly of alpha-synuclein leading to the accelerated amyloid fibril formation with shear stress. *PLoS One* **4,** e4177 (2009).

87. Kashchiev, D. *Nucleation: Basic Theory with Applications*. (Butterworth-Heinemann, 2000).

88. Kashchiev, D. & Auer, S. Nucleation of amyloid fibrils. *J. Chem. Phys.* **132,** 215101 (2010).

89. Zhang, T. H. & Liu, X. Y. Nucleation: what happens at the initial stage? *Angew. Chem. Int. Ed. Engl.* **48,** 1308–12 (2009).

90. Padrick, S. B. & Miranker, A. D. Islet amyloid: phase partitioning and secondary nucleation are central to the mechanism of fibrillogenesis. *Biochemistry* **41,** 4694–703 (2002).

91. Heldt, C. L., Zhang, S. & Belfort, G. Asymmetric amyloid fibril elongation: a new perspective on a symmetric world. *Proteins* **79,** 92–8 (2011).

92. Ban, T. & Goto, Y. Direct Observation of Amyloid Growth Monitored by Total Internal Reflection Fluorescence Microscopy. *Methods Enzymol.* **413,** 91–102 (2006).

93. Collins, S. R., Douglass, A., Vale, R. D. & Weissman, J. S. Mechanism of prion propagation: amyloid growth occurs by monomer addition. *PLoS Biol.* **2,** e321 (2004).

94. Patil, S. M., Mehta, A., Jha, S. & Alexandrescu, A. T. Heterogeneous amylin fibril growth mechanisms imaged by total internal reflection fluorescence microscopy. *Biochemistry* **50,** 2808–19 (2011).

95. Scheibel, T., Kowal, a S., Bloom, J. D. & Lindquist, S. L. Bidirectional amyloid fiber growth for a yeast prion determinant. *Curr. Biol.* **11,** 366–9 (2001).

96. Ionescu-Zanetti, C. *et al.* Monitoring the assembly of Ig light-chain amyloid fibrils by atomic force microscopy. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 13175–9 (1999).

97. Martins, P. M. & Rocha, F. New developments on size-dependent growth applied to the crystallization of sucrose. *Surf. Sci.* **601,** 5466–5472 (2007).

98. Martins, P. M. & Rocha, F. Characterization of crystal growth using a spiral nucleation model. *Surf. Sci.* **601,** 3400–3408 (2007).

99. Martins, P. M. & Rocha, F. a. Kinetic studies on the influence of temperature and growth rate history on crystal growth. *Cryst. Res. Technol.* **43,** 1258–1267 (2008).

100. Chatani, E. *et al.* Pre-steady-state kinetic analysis of the elongation of amyloid fibrils of beta(2)-microglobulin with tryptophan mutagenesis. *J. Mol. Biol.* **400,** 1057–66 (2010).

101. Konuma, T. *et al.* Kinetic intermediates of β(2)-microglobulin fibril elongation probed by pulse-labeling H/D exchange combined with NMR analysis. *J. Mol. Biol.* **405,** 851–62 (2011).

102. Santoso, a, Chien, P., Osherovich, L. Z. & Weissman, J. S. Molecular basis of a yeast prion species barrier. *Cell* **100,** 277–88 (2000).

103. Jiménez, J. L. *et al.* The protofilament structure of insulin amyloid fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 9196–201 (2002).

104. Smith, J. F., Knowles, T. P. J., Dobson, C. M., Macphee, C. E. & Welland, M. E. Characterization of the nanoscale properties of individual amyloid fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 15806–11 (2006).

105. Xue, W.-F. *et al.* Fibril fragmentation enhances amyloid cytotoxicity. *J. Biol. Chem.* **284,** 34272–82 (2009).

106. Ohhashi, Y., Kihara, M., Naiki, H. & Goto, Y. Ultrasonication-induced amyloid fibril formation of beta2-microglobulin. *J. Biol. Chem.* **280,** 32843–8 (2005).

107. Kim, J. *et al.* Pyrroloquinoline quinone inhibits the fibrillation of amyloid proteins. *Prion* **4,** 26–31 (2010).

108. Giehm, L. & Otzen, D. E. Strategies to increase the reproducibility of protein fibrillization in plate reader assays. *Anal. Biochem.* **400,** 270–81 (2010).

109. Fändrich, M. Absolute correlation between lag time and growth rate in the spontaneous formation of several amyloid-like aggregates and fibrils. *J. Mol. Biol.* **365,** 1266–70 (2007).

110. Sasahara, K., Yagi, H., Sakai, M., Naiki, H. & Goto, Y. Amyloid nucleation triggered by agitation of beta2-microglobulin under acidic and neutral pH conditions. *Biochemistry* **47,** 2650–60 (2008).

111. Rochet, J. C. & Lansbury, P. T. Amyloid fibrillogenesis: themes and variations. *Curr. Opin. Struct. Biol.* **10,** 60–8 (2000).

112. Radford, S. E., Gosal, W. S. & Platt, G. W. Towards an understanding of the structural molecular mechanism of beta(2)-microglobulin amyloid formation in vitro. *Biochim. Biophys. Acta* **1753,** 51–63 (2005).

113. Tanaka, M., Collins, S. R., Toyama, B. H. & Weissman, J. S. The physical basis of how prion conformations determine strain phenotypes. *Nature* **442,** 585–589 (2006).

114. White, D. a, Buell, A. K., Knowles, T. P. J., Welland, M. E. & Dobson, C. M. Protein aggregation in crowded environments. *J. Am. Chem. Soc.* **132,** 5170–5 (2010).

115. Baskakov, I. V & Bocharova, O. V. In Vitro Conversion of Mammalian Prion Protein into Amyloid Fibrils Displays Unusual Features. *Biochemistry* **44,** 2339–2348 (2005).

116. Pronchik, J., He, X., Giurleo, J. T. & Talaga, D. S. In vitro formation of amyloid from alpha-synuclein is dominated by reactions at hydrophobic interfaces. *J. Am. Chem. Soc.* **132,** 9797–803 (2010).

117. Toyama, B. H., Kelly, M. J. S., Gross, J. D. & Weissman, J. S. The structural basis of yeast prion strain variants. *Nature* **449,** 233–7 (2007).

118. Kulmala, M. *et al.* Direct Observations of Atmospheric Aerosol Nucleation. *Science (80-.).* **339,** 943–946 (2013).

119. Kollman, J. M., Merdes, A., Mourey, L. & Agard, D. a. Microtubule nucleation by γ-tubulin complexes. *Nat. Rev. Mol. Cell Biol.* **12,** 709–721 (2011).

120. Wang, Y., Azaïs, T., Robin, M., Vallée, A., Catania, C., Legriel, P., Pehau-Arnaudet, G., Babonneau, F., Giraud-Guille, M., and Nassif, N. The predominant role of collagen in the nucleation, growth, structure and orientation of bone apatite. *Nat. Mater.* **11,** 724–733 (2012).

121. Glabe, C. G. Structural classification of toxic amyloid oligomers. *J. Biol. Chem.* **283,** 29639–29643 (2008).

122. Uversky, V. N. & Fink, A. L. Conformational constraints for amyloid fibrillation: the importance of being unfolded. *Biochim. Biophys. Acta* **1698,** 131–53 (2004).

123. Krishnan, R. *et al.* Conserved features of intermediates in amyloid assembly determine their benign or toxic states. *Proc. Natl. Acad. Sci.* **109,** 11172–11177 (2012).

124. Bieschke, J., Herbst, M., Wiglenda, T., Friedrich, R. P., Boeddrich, A., Schiele, F., Kleckers, D. & del Amo, J. M. L., Grüning, B. A., and Wang, Q. Small-molecule conversion of toxic oligomers to nontoxic β-sheet–rich amyloid fibrils. *Nat. Chem. Biol.* **8,** 93–101 (2012).

125. Hopping, G. *et al.* Designed α-sheet peptides inhibit amyloid formation by targeting toxic oligomers. *Elife* **3,** 1–14 (2014).

126. Ehrnhoefer, D. E. *et al.* EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nat. Struct. Mol. Biol.* **15,** 558–566 (2008).

127. Lee, J., Culyba, E. K., Powers, E. T. & Kelly, J. W. Amyloid-β forms fibrils by nucleated conformational conversion of oligomers. **7,** 602–609 (2012).

128. Wetzel, R. Kinetics and thermodynamics of amyloid fibril assembly. *Acc. Chem. Res.* **39,** 671–679 (2006).

129. Ruggeri, F. S. *et al.* Infrared nanospectroscopy characterization of oligomeric and fibrillar aggregates during amyloid formation. *Nat. Commun.* **6,** 7831 (2015).

130. Martins, P. M. True and apparent inhibition of amyloid fibril formation. *Prion* **7,** 136–9 (2013).

131. Juárez, J., Taboada, P. & Mosquera, V. Existence of different structural intermediates on the fibrillation pathway of human serum albumin. *Biophys. J.* **96,** 2353–2370 (2009).

132. Holm, N. K. *et al.* Aggregation and fibrillation of bovine serum albumin. *Biochim. Biophys. Acta - Proteins Proteomics* **1774,** 1128–1138 (2007).

133. Faria, T. Q. *et al.* A look into amyloid formation by transthyretin: aggregation pathway and a novel kinetic model. *Phys. Chem. Chem. Phys.* **17,** 7255–63 (2015).

134. Chatani, E., Yagi, H., Naiki, H. & Goto, Y. Polymorphism of β2-microglobulin amyloid fibrils manifested by

ultrasonication-enhanced fibril formation in trifluoroethanol. *J. Biol. Chem.* **287,** 22827–37 (2012).

135. Ramachandran, G. & Udgaonkar, J. B. Understanding the Kinetic Roles of the Inducer Heparin and of Rod-like Protofibrils during Amyloid Fibril Formation by Tau Protein. *J. Biol. Chem.* **286,** 38948–38959 (2011).

136. Yang, S., Griffin, M. D. W., Binger, K. J., Schuck, P. & Howlett, G. J. An equilibrium model for linear and closed-loop amyloid fibril formation. *J. Mol. Biol.* **421,** 364–77 (2012).

137. O'Nuallain, B. *et al.* Amyloid β-Protein Dimers Rapidly Form Stable Synaptotoxic Protofibrils. *J. Neurosci.* **30,** 14411–14419 (2010).

138. Rangachari, V. *et al.* Amyloid-B(1-42) rapidly forms protofibrils and oligomers by distinct pathways in low concentrations of sodium dodecylsulfate. *Biochemistry* **46,** 12451–12462 (2007).

139. Gibson, T. J. & Murphy, R. M. Inhibition of insulin fibrillogenesis with targeted peptides. *Protein Sci.* **15,** 1133–1141 (2006).

140. F. Oosawa and M. Kasai. A theory of linear and helical aggregations of macromolecules. *J. Mol. Biol.* **4,** 10–21 (1962).

141. Cohen, S. I. a. *et al.* Proliferation of amyloid- 42 aggregates occurs through a secondary nucleation mechanism. *Proc. Natl. Acad. Sci.* **110,** 9758–9763 (2013).

142. Self-assembly, F. *et al.* Article Competition between Primary Nucleation and Autocatalysis in Amyloid. *Biophysj* **108,** 632–643 (2015).

143. Cohen, S. I. a. *et al.* Nucleated polymerization with secondary pathways. I. Time evolution of the principal moments. *J. Chem. Phys.* **135,** 065105 (2011).

144. Morris, R. J. *et al.* Mechanistic and environmental control of the prevalence and lifetime of amyloid oligomers. *Nat. Commun.* **4,** 1891 (2013).

145. Brender, J. R. *et al.* Probing the Sources of the Apparent Irreproducibility of Amyloid Formation: Drastic Changes in Kinetics and a Switch in Mechanism Due to Micellelike Oligomer Formation at Critical Concentrations of IAPP. *J. Phys. Chem. B* **119,** 2886–2896 (2015).

146. Ow, S.-Y. & Dunstan, D. E. The effect of concentration, temperature and stirring on hen egg white lysozyme amyloid formation. *Soft Matter* **9,** 9692 (2013).

147. Hill, S. E., Miti, T., Richmond, T. & Muschol, M. Spatial Extent of Charge Repulsion Regulates Assembly Pathways for Lysozyme Amyloid Fibrils. *PLoS One* **6,** e18171 (2011).

148. Mishra, R. *et al.* Lysozyme amyloidogenesis is accelerated by specific nicking and fragmentation but decelerated by intact protein binding and conversion. *J. Mol. Biol.* **366,** 1029–44 (2007).

149. Umemoto, A., Yagi, H., So, M. & Goto, Y. High-throughput Analysis of Ultrasonication-forced Amyloid Fibrillation Reveals the Mechanism Underlying the Large Fluctuation in the Lag Time. *J. Biol. Chem.* **289,** 27290–9 (2014).

150. Szavits-Nossan, J., Eden, K., Morris, R. J., MacPhee, C. E., Evans, M. R., and Allen, R. J. Inherent variability in the

kinetics of autocatalytic protein self-assembly. *Phys. Rev. Lett.* **113,** 098–101 (2014).

151. Breydo, L. *et al.* The crowd you're in with: Effects of different types of crowding agents on protein aggregation. *Biochim. Biophys. Acta - Proteins Proteomics* **1844,** 346–357 (2014).

152. Morshedi, D., Rezaei-Ghaleh, N., Ebrahim-Habibi, A., Ahmadian, S. & Nemat-Gorgani, M. Inhibition of amyloid fibrillation of lysozyme by indole derivatives--possible mechanism of action. *FEBS J.* **274,** 6415–25 (2007).

153. Streets, A. M., Sourigues, Y., Kopito, R. R., Melki, R. & Quake, S. R. Simultaneous Measurement of Amyloid Fibril Formation by Dynamic Light Scattering and Fluorescence Reveals Complex Aggregation Kinetics. *PLoS One* **8,** e54541 (2013).

154. Mittal, S., Chowhan, R. K. & Singh, L. R. Macromolecular crowding: Macromolecules friend or foe. *Biochim. Biophys. Acta - Gen. Subj.* **1850,** 1822–1831 (2015).

155. Munishkina, L. a., Ahmad, A., Fink, A. L. & Uversky, V. N. Guiding protein aggregation with macromolecular crowding. *Biochemistry* **47,** 8993–9006 (2008).

156. Kuznetsova, I. M., Turoverov, K. K. & Uversky, V. N. *What macromolecular crowding can do to a protein. International journal of molecular sciences* **15,** (2014).

157. Hall, D. & Hoshino, M. Effects of macromolecular crowding on intracellular diffusion from a single particle perspective. *Biophys. Rev.* **2,** 39–53 (2010).

158. Lee, C. F., Bird, S., Shaw, M., Jean, L. & Vaux, D. J. Combined Effects of Agitation, Macromolecular

Crowding, and Interfaces on Amyloidogenesis. *J. Biol. Chem.* **287,** 38006–38019 (2012).

159. Benton, L. a, Smith, A. E., Young, G. B. & Pielak, G. J. Unexpected Effects of Macromolecular Crowding on Protein Stability Unexpected Effects of Macromolecular Crowding on Protein Stability. 9773–9775 (2012). doi:10.1021/bi300909q

160. Sarkar, M., Li, C. & Pielak, G. J. Soft interactions and crowding. *Biophys. Rev.* **5,** 187–194 (2013).

161. Uversky, V. N. Mysterious oligomerization of the amyloidogenic proteins. *FEBS J.* **277,** 2940–2953 (2010).

162. Lee, J., Culyba, E. K., Powers, E. T. & Kelly, J. W. Amyloid-β forms fibrils by nucleated conformational conversion of oligomers. *Nat. Chem. Biol.* **7,** 602–9 (2011).

163. Kaganovich, D., Kopito, R. & Frydman, J. {M}isfolded proteins partition between two distinct quality control compartments. *Nature* **454,** 1088–1095 (2008).

164. Hyung, S.-J. *et al.* Insights into antiamyloidogenic properties of the green tea extract (-)-epigallocatechin-3-gallate toward metal-associated amyloid-β species. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 3743–8 (2013).

165. Ladiwala, A. R. a *et al.* Conformational differences between two amyloid ??oligomers of similar size and dissimilar toxicity. *J. Biol. Chem.* **287,** 24765–24773 (2012).

166. Deva, T. *et al.* Off-pathway aggregation can inhibit fibrillation at high protein concentrations. *Biochim. Biophys. Acta - Proteins Proteomics* **1834,** 677–687 (2013).

167. Crespo, R. *et al.* What can the kinetics of amyloid fibril formation tell about off-pathway aggregation? *J. Biol. Chem.* jbc.M115.699348 (2015). doi:10.1074/jbc.M115.699348

168. Kodali, R. & Wetzel, R. Polymorphism in the intermediates and products of amyloid assembly. *Curr. Opin. Struct. Biol.* **17,** 48–57 (2007).

169. Ellis, R. J. Macromolecular crowding: Obvious but underappreciated. *Trends Biochem. Sci.* **26,** 597–604 (2001).

170. Hall, D. & Minton, A. P. Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges. *Biochim. Biophys. Acta - Proteins Proteomics* **1649,** 127–139 (2003).

171. Elcock, A. H. Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. *Curr. Opin. Struct. Biol.* **20,** 196–206 (2010).

172. Knowles, TP, M Vendruscolo, and C. D. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15,** 384–396 (2014).

173. Lo Nostro, P. & Ninham, B. W. Hofmeister phenomena: an update on ion specifLo Nostro, P., & Ninham, B. W. (2012). Hofmeister phenomena: an update on ion specificity in biology. Chemical Reviews, 112(4), 2286–322. doi:10.1021/cr200271jicity in biology. *Chem. Rev.* **112,** 2286–322 (2012).

174. Souillac, P. O. *et al.* Elucidation of the molecular mechanism during the early events in immunoglobulin light chain amyloid fibrillation. Evidence for an off-

pathway oligomer at acidic pH. *J. Biol. Chem.* **277,** 12666–12679 (2002).

175. Stine, WB, L Jungbauer, C Yu, and M. L. *Alzheimer's Disease and Frontotemporal Dementia*. (Humana Press., 2011).

176. Hellstrand, E., Boland, B., Walsh, D. M. & Linse, S. Amyloid ??-protein aggregation produces highly reproducible kinetic data and occurs by a two-phase process. *ACS Chem. Neurosci.* **1,** 13–18 (2010).

177. Munishkina, L. A., Henriques, J., Uversky, V. N. & Fink, A. L. Role of protein-water interactions and electrostatics in alpha-synuclein fibril formation. *Biochemistry* **43,** 3289–3300 (2004).

178. Klement, K. *et al.* Effect of different salt ions on the propensity of aggregation and on the structure of Alzheimer's abeta(1-40) amyloid fibrils. *J. Mol. Biol.* **373,** 1321–33 (2007).

179. Sharp, K. a. Analysis of the size dependence of macromolecular crowding shows that smaller is better. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 7990–5 (2015).

180. Politou, A. & Temussi, P. A. Revisiting a dogma: the effect of volume exclusion in molecular crowding. *Curr. Opin. Struct. Biol.* **30,** 1–6 (2015).

181. Sapir, L. & Harries, D. Is the depletion force entropic? Molecular crowding beyond steric interactions. *Curr. Opin. Colloid Interface Sci.* **20,** 3–10 (2015).

182. Pastore, A. & Temussi, P. A. The two faces of Janus: Functional interactions and protein aggregation. *Curr. Opin. Struct. Biol.* **22,** 30–37 (2012).

183. Sabate, R., De Groot, N. S. & Ventura, S. Protein folding and aggregation in bacteria. *Cell. Mol. Life Sci.* **67,** 2695–2715 (2010).

184. Lashuel, H. a, Overk, C. R., Oueslati, A. & Masliah, E. The many faces of α-synuclein: from structure and toxicity to therapeutic target. *Nat. Rev. Neurosci.* **14,** 38–48 (2013).

185. Outeiro, T. F. & Kazantsev, A. Drug Targeting of alpha-Synuclein Oligomerization in Synucleinopathies. *Perspect. Medicin. Chem.* **2,** 41–49 (2008).

186. Schreurs, S. *et al.* In vitro phosphorylation does not influence the aggregation kinetics of WT α-synuclein in contrast to its phosphorylation mutants. *Int. J. Mol. Sci.* **15,** 1040–1067 (2014).

187. Coelho-Cerqueira, E., Carmo-Gonçalves, P., Sá Pinheiro, A., Cortines, J. & Follmer, C. α-Synuclein as an intrinsically disordered monomer - Fact or artefact? *FEBS J.* **280,** 4915–4927 (2013).

188. Öhrfelt, A. *et al.* Cerebrospinal fluid α-synuclein in neurodegenerative disorders-A marker of synapse loss? *Neurosci. Lett.* **450,** 332–335 (2009).

189. Brown, D. R. α-Synuclein as a ferrireductase. *Biochem. Soc. Trans.* **41,** 1513–7 (2013).

190. Gibrat, C. *et al.* Differences between subacute and chronic MPTP mice models: Investigation of dopaminergic neuronal degeneration and α-synuclein inclusions. *J. Neurochem.* **109,** 1469–1482 (2009).

191. Williams, M. Productivity Shortfalls in Drug Discovery : Contributions from the Preclinical Sciences ? **336,** 3–8 (2011).

192. Kubinyi, H. Drug research: myths, hype and reality. **2,** 5–8 (2003).

193. Apetri, M. M., Maiti, N. C., Zagorski, M. G., Carey, P. R. & Anderson, V. E. Secondary structure of α-synuclein oligomers: Characterization by Raman and atomic force microscopy. *J. Mol. Biol.* **355,** 63–71 (2006).

194. Loureiro, J. a. *et al.* Fluorinated beta-sheet breaker peptides. *J. Mater. Chem. B* **2,** 2259 (2014).

195. Roostaee, A., Beaudoin, S., Staskevicius, A. & Roucou, X. Aggregation and neurotoxicity of recombinant alpha-synuclein aggregates initiated by dimerization. *Mol. Neurodegener.* **8,** 5 (2013).

196. El-Agnaf, O. M. a. & Irvine, G. B. Aggregation and properties of α–synuclein and related proteins. *Spectroscopy* **15,** 141–150 (2001).

197. Zhu, M. *et al.* The flavonoid baicalein inhibits fibrillation of α-synuclein and disaggregates existing fibrils. *J. Biol. Chem.* **279,** 26846–26857 (2004).

198. So, M. *et al.* Ultrasonication-dependent acceleration of amyloid fibril formation. *J. Mol. Biol.* **412,** 568–577 (2011).